An Oscillation Theory of Handwriting

by

John M. Hollerbach

Massachusetts Institute of Technology

March, 1980

This report is a revised version of a dissertation submitted to the Department of Electrical Engineering and Computer Science on August 11, 1978 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Abstract

Handwriting production is viewed as a constrained modulation of an underlying oscillatory process. Coupled oscillations in horizontal and vertical directions produce letter forms, and when superimposed on a rightward constant velocity horizontal sweep result in spatially separated letters. Modulation of the vertical oscillation is responsible for control of letter height, either through altering the frequency or altering the acceleration amplitude. Modulation of the horizontal oscillation is responsible for control of corner shape through altering phase or amplitude.

The vertical velocity zero crossing in the velocity space diagram is important from the standpoint of control. Changing the horizontal velocity value at this zero crossing controls corner shape, and such changes can be effected through modifying the horizontal oscillation amplitude and phase. Changing the slope at this zero crossing controls writing slant: this slope depends on the horizontal and vertical velocity amplitudes and on the relative phase difference. Letter height modulation is also best applied at the vertical velocity zero crossing to preserve an even baseline. The corner shape and slant constraints completely determine the amplitude and phase relations between the two oscillations. Under these constraints interletter separation is not an independent parameter.

This theory applies generally to a number of acceleration oscillation patterns such as sinusoidal, rectangular and trapezoidal oscillations. The oscillation theory also provides an explanation for how handwriting might degenerate with speed.

An implementation of the theory in the context of the spring muscle model is developed. Here sinusoidal oscillations arise from a purely mechanical sources; orthogonal antagonistic spring pairs generate particular cycloids depending on the initial conditions. Modulating betwen cycloids can be achieved by changing the spring zero settings at the appropriate times. Frequency can be modulated either by shifting between coactivation and alternating activation of the antagonistic springs or by presuming variable spring constant springs.

An acceleration and position measuring apparatus was developed for measurements of human handwriting. Measurements of human writing are consistent with the oscillation theory.

It is shown that the minimum energy movement for the spring muscle model is bang-coast-bang. For certain parameter values a singular arc solution can be shown to be minimizing. Experimental measurements however indicate that handwriting is not a minimum energy movement.

Acknowledgments

I owe a large debt of gratitude to David Marr for supervising this research and for his infectious enthusiasm. I would also like to thank my readers Emilio Bizzi for invaluable advice in later stages of the thesis, Berthold Horn for incisive comments, and Murray Eden for perspective on the research.

A number of people made significant contributions to the success of this effort. Mike Bowles introduced me to the techniques of modern control theory and pointed out its applicability to this research. Fred Drenkham played a key role in building the electronics to the accelerometer apparatus and in maintaining the VICARM. John Purbrick built the pen attachment to the VICARM. The expert craftsmanship that went into the construction for the accelerometer apparatus by Ignacio Garabieta made accurate measurement feasible. Marc Raibert and Mike Brady provided constructive criticisms.

Lastly, I wish to thank Patrick Winston for arranging funding for this research. To all the members of the AI lab, and especially to those who have delighted in the grape with me, I extend my gratitude for creating a vital and high calibre atmosphere that greatly facilitated research.

Table of Contents

Chapter 1. Introduction

This report addresses the motor control of handwriting. The goal has been to arrive at a detailed understanding of the motor programs underlying the production of diverse letter shapes. A theory of handwriting production is presented that views handwriting as a constrained modulation of an underlying oscillatory process. Coupled oscillations in horizontal and vertical directions produce letter forms, and when superimposed on a rightward constant velocity horizontal sweep result in spatially separated letters. Modulation of the vertical oscillation is responsible for control of letter height, either through altering the frequency or altering the acceleration amplitude. Modulation of the horizontal oscillation is responsible for control of corner shape through altering phase or amplitude. The horizontal sweep velocity is presumed not to modulate.

In understanding handwriting control the velocity space diagram, which is a plot of horizontal velocity versus vertical velocity, is an important tool. One point in particular, the vertical velocity zero crossing, seems important from the standpoint of control. Changing the horizontal velocity value at this zero crossing controls corner shape, and such changes can be effected through modifying the horizontal oscillation amplitude and phase.

Writing slant can be explained in this theory as an artifact of coupled orthogonal oscillations, and is dependent on the horizontal and vertical velocity amplitudes and the relative phase difference. The slant angle is given by the tangent to the velocity space diagram at the vertical velocity zero crossing. Merely by altering the slope and intercept at this zero crossing therefore one obtains control of two of the important features of letter shape.

Since modulating letter height or corner shape ordinarily results in velocity amplitude and phase changes, an applied modulation can be expected to change writing slant. A consistent writing slant is presumably a stylistic constraint in handwriting, and it would not be considered acceptable writing if for example tall letters were slanted randomly or if tall letters like *l*'s were slanted differently from short letters like *e*'s. Thus there are constraints on modulations to preserve writing slant. Together with corner shape constraints, for example

making sure an *l* has a loop top corner of the appropriate fullness, the slant constraint completely determines the amplitude and phase relations between the two oscillations. Under these constraints interletter separation is not an independent parameter. Letter height modulation is best applied at zero vertical velocity to preserve an even baseline. If applied at the bottom corner upper zone letters such as *l* and *h* can be produced; if applied at the top corner lower zone letters such as *g* and *y* can be produced. The writing slant constraint suggests that frequency modulation is preferable to acceleration amplitude modulation of the vertical oscillation for letter height. The slant measure is independent of frequency, and if letter height can be modulated through frequency alone then there would not be any problem in preserving writing slant. Unfortunately the corner shape constraint usually requires a modulation of the horizontal oscillation amplitude and phase as well as of the frequency during letter height modulation, but in some sense a frequency modulation for letter height is easier to control than an amplitude modulation for letter height. Experiments on human handwriting show that people use a mixture of frequency and amplitude modulation for letter height, but that the frequency modulation is the dominant contributor to letter height.

This theory, which applies generally to a number of acceleration oscillation patterns such as rectangular and trapezoidal oscillations, has been particularized to sinusoids for mathematical convenience. Sinusoidal based oscillations produce writing that belongs to the class of cycloidal curves. The process of modulation can be considered as one of shifting between different cycloids.

To explain one way in which the human motor system might implement this theory a spring muscle model is invoked. This model likens the action of muscle to a spring with variable zero setting. The amount of force generated by the muscle can be changed merely by controlling the zero setting, which is identified with changing the firing rate of the α-motorneurons. With this spring muscle model a sinusoidal oscillation arises from a purely mechanical source, since antagonistic springs form a harmonic oscillator. Orthogonal antagonistic spring pairs generate particular cycloids depending on the initial conditions. Modulating betwen cycloids can be achieved by changing the spring zero settings at the appropriate times.

Frequency modulation in a fixed spring constant model can only be achieved in a restricted way. By

shifting between coactivation and alternating activation of the antagonistic springs a frequency modulation proportional to $\sqrt{2}$ can be obtained. For more flexible frequency modulation one has to presume a variable stiffness spring, with the zero setting changing not only the amount of applied force but also the spring constant. Under a variable stiffness spring model it is possible to modulate frequency arbitrarily and to satisfy as well the constraints imposed by slant constancy and corner shaping. Measurements on human subjects indicate a frequency modulation in the broad neighborhood of $\sqrt{2}$ in going between tall and short letters.

The oscillation theory was tested by means of a special apparatus for measuring handwriting of human subjects. An x-y sliding rail system with 6 degrees of freedom was mounted with accelerometers, and a special holder incorporated a writing tablet pen. Acceleration and position are direct measurements with this apparatus, and velocity is obtained by processing one of these signals.

Measurements of human writing are consistent with the oscillation theory. The assumption of a constant velocity horizontal sweep is substantiated by the data. Subjects demonstrate slant constancy and corner shape control during letter height modulation as evidenced in their velocity space diagrams: the vertical velocity zero crossings remain constant both in intercept and in slope. The underlying oscillation pattern seen in human subjects is somewhat equivocal; some appear sinusoidal, others are trapezoidal, and a few are almost rectangular.

The oscillation theory provides an explanation for how handwriting might degenerate with speed. If one presumes that the difficulty in producing a particular letter shape is related to the number and the severity of phase, frequency, and amplitude modulations, then it can be predicted that difficult modulations would be abbreviated and that the resultant writing would be simplified. The one modulation which yields the greatest difficulty is a transition between clockwise and counterclockwise movement, such as between an *e* and and *n* or within a letter such as an *h* or a *y*. I have observed that fast writers eliminate all clockwise movements. Rather than sending fast writers to rural communes for rehabilitation, the culprit seems to be the Palmer script itself which is not tailored for speed. It is possible to design a cursive script that eliminates clockwise movement and which does not degenerate with speed.

In a separate chapter it is shown that the minimum energy movement for the spring muscle model is bang-coast-bang. For certain parameter values a singular arc solution can be shown to be minimizing. If handwriting were a minimum energy movement, one would expect a roughly trapezoidal oscillation pattern. The acceleration amplitudes would always be at maximum, and letter height modulation would be achieved solely through frequency modulation. Small handwriting would also be executed faster than large handwriting. Unfortunately experimental measurements do not bear out these expectations. The speed of handwriting is independent of writing size. Amplitude modulation for letter height occurs as well as frequency modulation. One can conclude that handwriting is not a minimum energy movement.

Handwriting and Biomechanical Complexity

The oscillation theory is a description of handwriting trajectories and of the constraints that apply in their formation. It says nothing about how the biological agent arranges its joints, limbs, and muscles to satisfy the trajectory constraints. Although this report does not address this question it is nevertheless useful to consider what problems the agent must solve to implement this theory. The agent is required by the theory to provide three functional degrees of freedom, one each for the orthogonal horizontal and vertical oscillations and a third for the horizontal sweep. A functional degree of freedom may or may not conform to a particular joint; it could arise from a synergistic action of a number of muscles and joints. Human subjects show a bewildering variety of muscle and joint usage in handwriting; the particular choice of mechanical arrangement is not important as long as the arrangement can satisfy the constraints of the theory. In a sense the great mechanical complexity of the hand serves not to complicate handwriting but to provide a variety of configurations.

It is also up to the agent to ensure that the motion remains planar and that the writing implement remains on the writing surface with a particular pressure. Some adjustment of the functional synergy is therefore required depending on the range of joint extension. A further complexity is the requirement of holding the writing implement. The holding has to adapt to changing joint angles: human subjects find it necessary to roll the writing implement between fingers and thumb for this purpose. Lastly, gross positioning movements

between words, between lines, and for dotting i's and crossing t's are required but which lie outside the scope of the theory.

Related Work

It is significant that an evolutionarily advanced system, the human hand, seems to use a philogenetically old method of motor control for handwriting. The notion of oscillators and of modulations to the oscillators ties handwriting to locomotion [Grillner 1975]. The degree to which the handwriting oscillators arise from a mechanical source or from active programming needs to be determined before one can ask about the location of the oscillator; for example, whether the oscillators share a spinal cord or brainstem location with the more primitive locomotion oscillators.

With regard to past handwriting research, the focus has been more descriptive than it has been generative. Initial efforts were aimed at devising a measurement apparatus, such as spark marking of teledelthos paper with an iron stylus [Denier van der Gon and Thuring 1965] and an electrolytic water tank [McDonald 1966 and Yasuhara 1975]. Yasuhara [1975] attempted to interpret EMG measurements as force data. Herbst and Liu [1977] devised a pen with accelerometer mountings. Crane and Savoie [1977] mounted strain gauges onto a membrane which was distorted by a pen shaft during writing.

The approach to modelling a handwriting trajectory has usually been a curve fitting to measured or inferred accelerations. Mermelstein and Eden [1964] segmented writing for fitting with quarter sine waves. Denier van der Gon and Thuring [1965] assumed a rectangular form to the accelerations. McDonald [1966] fit trapezoids to the accelerations. Yasuhara [1975] assumed an exponential rise and decay time to an acceleration plateau. The end result of this process is a list of acceleration burst durations and amplitudes which when applied to the corresponding model yields synthetic writing close to the measured human handwriting.

Given the great biomechanical complexity of joints and muscles, any simple exercise in exact duplication of a particular human handwritten word would seem futile. It is also unclear that curve fitting with successively more baroque models adds more enlightenment to how the human motor system accomplishes handwriting.

Chapter 2. An Oscillation Theory of Handwriting

There are many strategies by which accelerations can be fashioned to produce acceptable handwriting. A particularly parsimonious view is that handwriting is produced by orthogonal oscillations horizontal and vertical in the plane of the writing surface, and that these two oscillations are superimposed on a constant velocity rightward horizontal sweep. The orthogonal oscillations are responsible for producing letter shape, while the horizontal sweep strings these letters out into a rightward moving train. The oscillations are modulated in certain ways and at specific points to produce the shapes characteristic of the English Palmer script.

It is mathematically convenient to model the oscillations by sinusoids, although the main conclusions presented in this chapter hold for other oscillation patterns such as trapezoids and rectangles. Another reason for selecting sinusoids is that in Chapter 3 it will be seen that a spring muscle model leads to a harmonic oscillator.

2.1 Cycloidal Writing

The equations governing the oscillations in the velocity domain can be written as:

$$\begin{aligned} \dot{x} &= a\sin(\omega_x(t-t_0)+\phi_x)+c \\ \dot{y} &= b\sin(\omega_y(t-t_0)+\phi_y) \end{aligned} \tag{2.1}$$

where a and b are the horizontal and vertical velocity amplitudes, ω_x and ω_y are the horizontal and vertical frequencies, ϕ_x and ϕ_y are the horizontal and vertical phases, t is the time with respect to the reference time t_0, and c is the magnitude of the horizontal sweep.

These velocity equations when integrated yield a *cycloid* [Lawrence 1972]. Consider for the moment the case where $a = b = c$, $\omega_x = \omega_y$, and $\phi_x - \phi_y = \pi/2$. The resulting cycloid can be thought of as the curve traced by a point on a rim of a disc rolling on a plane. When $c < a$ the tracing point is on a radial extension to the disc and the curve is called a *prolate cycloid*. When $c > a$ the tracing point is inside the rim and yields a

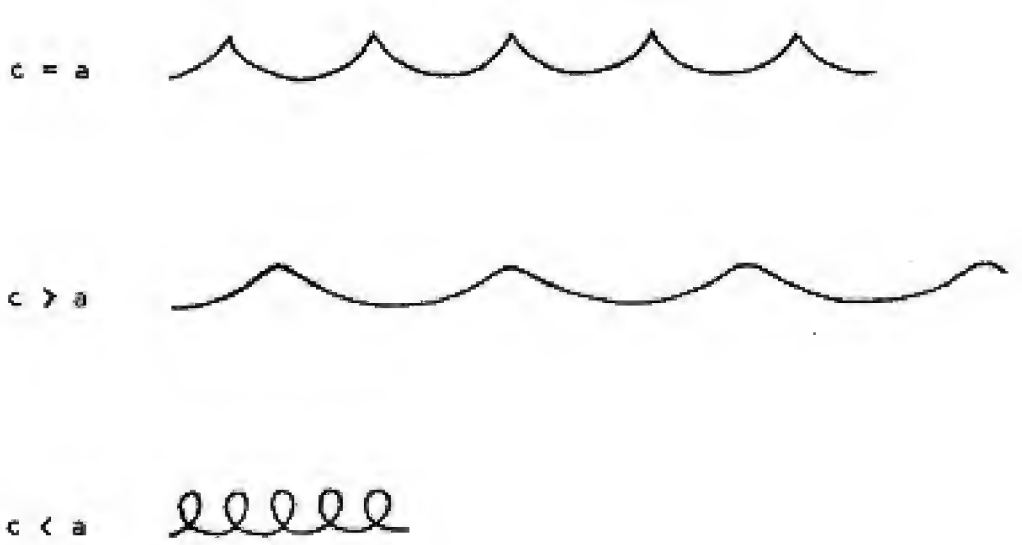


Figure 2.1. The top curve represents a cycloid, the middle curve a curtate cycloid, and the bottom curve a prolate cycloid.

curtate cycloid. Representative curves of each type are shown in Figure 2.1. We already see in these curves the inklings of basic handwriting patterns. An ordinary cycloid yields strokes with a sharp top corner, such as is found in *t*, *u*, and *w*. A prolate cycloid yields top loops as in the letter *e*. A curtate cycloid yields rounded top corners which are related to rounded corners in *h*, *m*, and *n*, but which are not fully apparent in this Figure.

The velocity equations represent an ellipse centered at $(c, 0)$. The tilt α and eccentricity e are derived below. The parametric velocity equations 2.1 become when t is eliminated:

$$b^2x^2 - 2ab\cos\phi\, xy + a^2y^2 - a^2b^2\sin^2\phi = 0$$

It is shown in [Thomas] that for a second degree polynomial $Ax^2 + Bxy + Cy^2 + F = 0$ the xy term results from a rotation of an ellipse. The amount of rotation α is given by $\cot 2\alpha = (C - A)/B$, or:

$$\cot 2\alpha = \frac{b^2 - a^2}{2ab\cos\phi} \tag{2.2}$$

It is further shown in [Thomas] that a rotation to set the xy term to zero gives an ellipse $A'x'^2 + C'y'^2 + F' = 0$ whose coefficients satisfy the relations $A' + C' = A + C$ and $-4A'C' = B^2 - 4AC$. These two relations may be used to solve for the new coefficients A' and C'.

$$A' = \frac{1}{2}(a^2 + b^2 + Q)$$
$$C' = \frac{1}{2}(a^2 + b^2 - Q)$$

where $Q = \sqrt{(a^2 + b^2)^2 - 4a^2b^2\sin^2\phi}$. The eccentricity e is given by $\frac{\sqrt{A'-C'}}{A'}$, or:

$$e = \sqrt{\frac{Q}{b^2 + a^2 + Q}} \tag{2.3}$$

The effect of setting $a \neq b$ is to turn the velocity space circle into an ellipse. The phase angle ϕ changes both the eccentricity and the tilt of this ellipse. In position space the sinusoidal portion of the integrated equations (2.1) represents an ellipse with the same eccentricity and tilt as the velocity space ellipse, but with a phase lag of 90 degrees.

The effect of phase shift on the velocity space ellipse is illustrated in the sequence in Figure 2.2. As the phase shifts from 90° to −30°, the ellipse changes its tilt and eccentricity. The ellipse corresponding to $\phi = -30°$ has a clockwise rotation sense, the similar shaped ellipse with $\phi = 30°$ a counterclockwise sense. Switching between counterclockwise and clockwise movement reverses the sense of top and bottom corners. The sharp top corners of letters such as u and i become the sharp bottom corners such as m and n.

The importance of this last Figure is that it shows that merely by setting up different initial conditions a train of the basic letter shapes in handwriting is produced. These basic shapes can be categorized as loop, sharp, and rounded top corners. By modulating the oscillation at specific times in the cycle and with specific phase and amplitude changes the oscillation train can be transformed from one basic pattern of shapes into another basic pattern. For example, Figure 2.2 indicates that if the phase could be altered one could obtain a

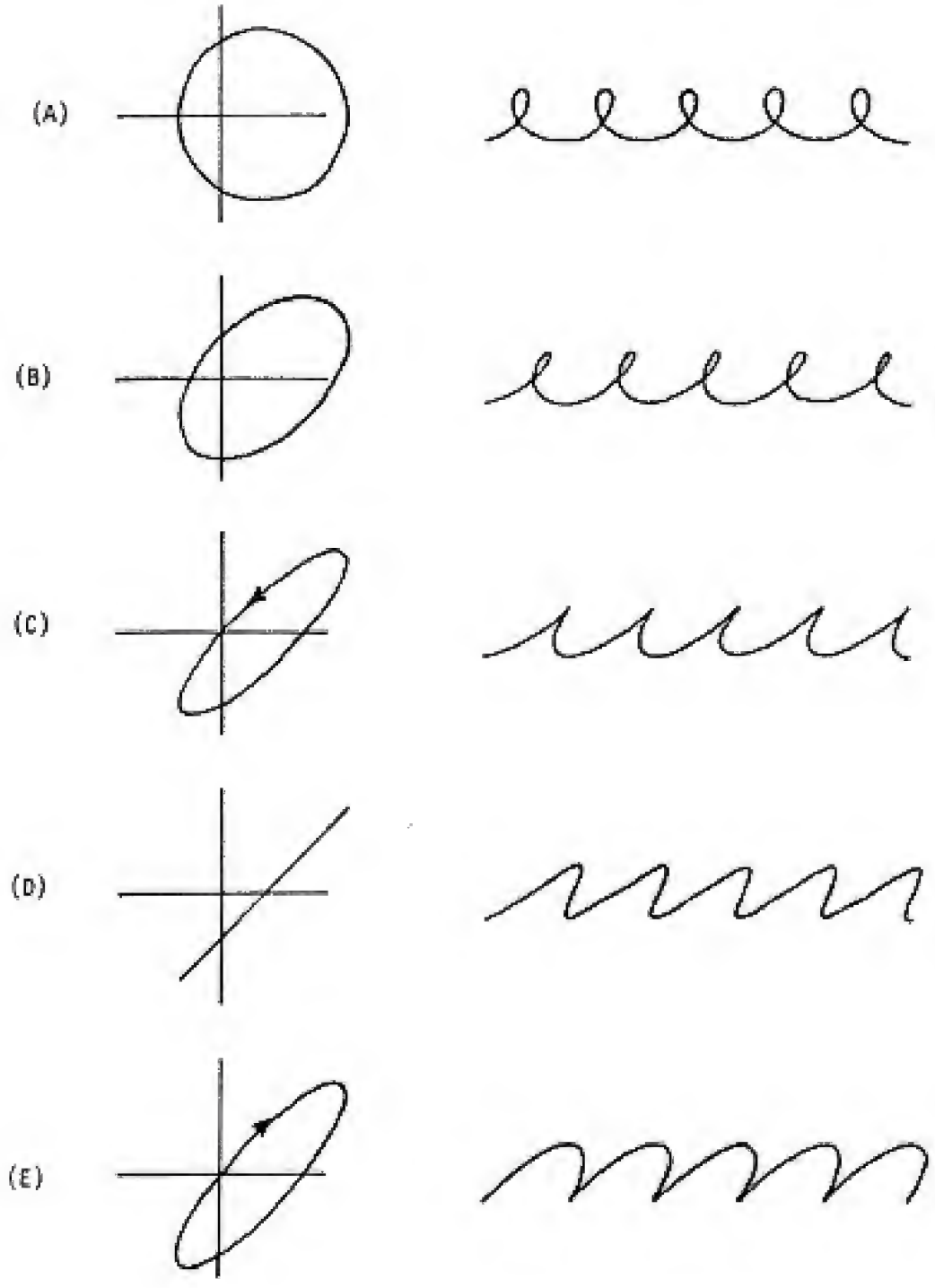


Figure 2.2. Phase modulation of a cycloid with parameters $a = b = 2e$. The phase shifts are: (a) 90 degrees, (b) 60 degrees, (c) 30 degrees, (d) 0 degrees, and (e) -30 degrees.

tailored sequence of top corner shapes. It is also clear that by modulating the vertical velocity amplitude letter height could be controlled.

Control of the vertical velocity zero crossing

The important parameter influencing letter shape is not actually phase change but the value of the horizontal velocity at the vertical velocity zero crossing. The vertical velocity is zero at the top corner when $\omega(t - t_0) + \phi_y = (2n + 1)\pi$. The value of the horizontal velocity at this zero crossing is:

$$\dot{x} = c - a \sin \phi \tag{2.4}$$

When this $\dot{x}$ is zero a sharp top corner results. The more negative this value, the fuller the top loop. The more positive this value, the rounder the top corner. A clockwise movement has the effect of reversing the sense of top and bottom corner. To illustrate that the vertical velocity zero crossing is primarily responsible for corner shape, the sequence of *e*'s in Figure 2.3 all have the same zero crossing but differ in velocity amplitude and phase shift. It can be seen that except for letter height the top corner shape is maintained.

To transform one cycloidal pattern into another cycloidal pattern, phase and amplitude modulations of the horizontal oscillation have to be applied at appropriate times and at the appropriate levels. It is conceivable that a modulation of the vertical oscillation could assist the shape transformation such as through a phase change, but there would be added complications of keeping an even baseline and of regulating letter height. For shape modulation in which letter height does not change, therefore, it is presumed that horizontal oscillation modulation alone occurs.

As an example of shape control by altering the vertical velocity zero crossing, an *e* cycloid in Figure 2.4 has been modulated to give the sequence *eune*. The parameters of the modulations to the horizontal oscillation are presented in Table 2.1. Frequency was set at 5 Hz; the vertical velocity amplitude was 28π mm/sec. The units are of course arbitrary, but they have a rough physiological correspondence.

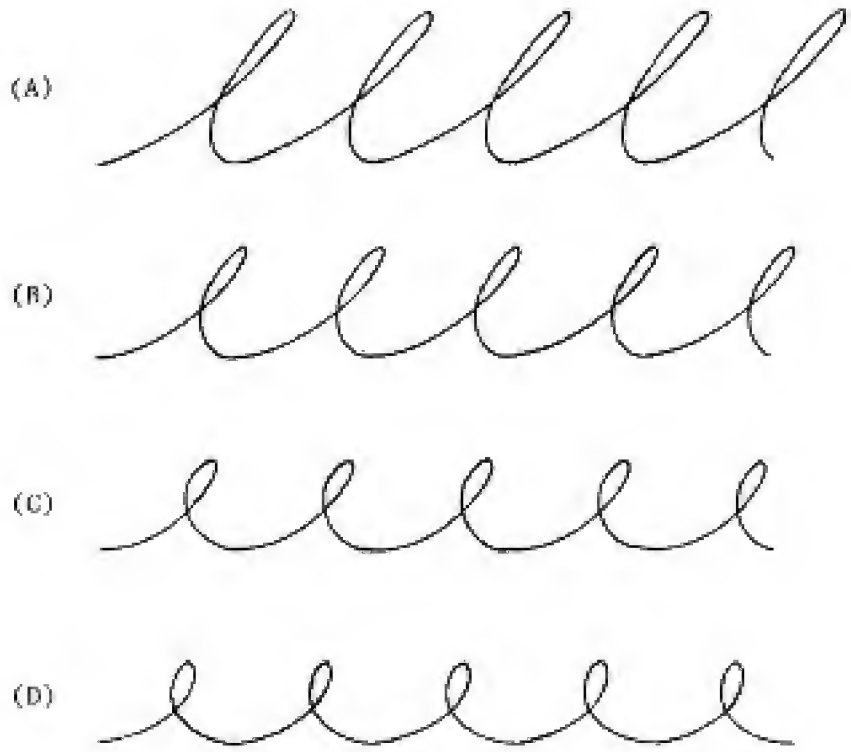


Figure 2.3. Cycloids generated with $a = 6$, $c = 20$, and with a constant vertical velocity zero crossing at $c - a \sin \phi = -14.64$, but different phase relations, indicate a similar top corner shape. This top corner shape is primarily controlled by the zero crossing instead of the phase difference. (A) $\phi = 30°$, $a = 69.28$; (B) $\phi = 45°$, $a = 48.99$; (C) $\phi = 60°$, $a = 40$; (D) $\phi = 75°$, $a = 35.16$.

Table 2.1

Time	ϕ	a	$c - a \sin \phi$
0.00 secs	43.66 degrees	65.16 mm/sec	-25.00 mm/sec
0.23 secs	17.03 degrees	68.28 mm/sec	0.00 mm/sec
0.64 secs	-7.50 degrees	75.64 mm/sec	29.87 mm/sec
0.95 secs	30.94 degrees	87.46 mm/sec	-25.00 mm/sec

To devise the modulations, a particular value for the vertical velocity zero crossing (2.4), which is the last column in Table 2.1, was chosen for producing a particular corner type. Arbitrary values for a and ϕ were then

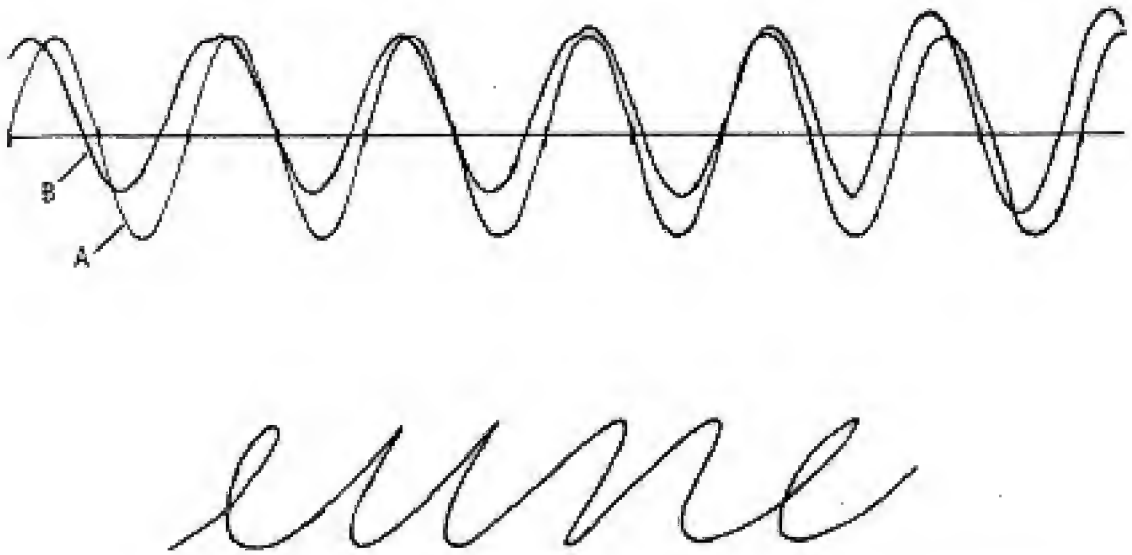


Figure 2.4. Modulation of the x oscillation by adjusting the vertical velocity zero crossing produces this sequence *eurse*. Above the writing, curve A is the vertical velocity, curve B the horizontal velocity.

chosen consistent with the desired zero crossing value. Examining the last column in Table 2.1, the effects of the various modulations on shape can be seen. The first modulation at 0.23 seconds results in a zero crossing value of 0.00 mm/sec; hence a sharp top corner is obtained. The second modulation gives a positive zero crossing and a rounded top corner. Finally the original *e* cycloid is obtained through a modulation that yields a zero crossing value of -25.00 mm/sec.

Modulation of Letter Height

Letter height may be modulated by changing either the vertical acceleration amplitude or the frequency of oscillation. To preserve an even baseline, the modulation is best applied at points of zero vertical velocity, which occur at the top and bottom corners. If the point of modulation is a bottom corner, tall letters such as *l* and *h* can be produced. If the point of modulation is a top corner, lower zone letters such as *g* and *y* can be produced. An example of amplitude modulation to achieve different letter heights is illustrated by the

Figure 2.5. Modulation of the vertical oscillation at points of zero vertical velocity yields different letter heights by a process of amplitude modulation without phase change.

alternating patterns of *el* in Figure 2.5. Since the horizontal oscillation is unchanged and since the vertical phase does not change because the modulation was applied at a point of zero vertical velocity, corner shape has been preserved during this height modulation. The velocity space ellipses corresponding to the *e*'s and *l*'s in Figure 2.5 are seen to have the same zero crossings.

2.2 Control of Writing Slant

A salient feature of Figure 2.5 is the difference in writing slant between the *e*'s and *l*'s, indicating that writing slant is somehow a function of the various parameters. The horizontal and vertical oscillations act orthogonally, yet slanted writing results from their combination.

To investigate the origin of writing slant, a psychophysical experiment was performed to obtain a measure for writing slant. A group of subjects were asked to estimate writing slant in an assortment of top cusp cycloids. The subjects chose with a high degree of agreement one of two strategies in estimating slant depending on the interpretation of the cycloid. If the cycloid was interpreted as a chain of *u*'s, the slant was taken as the line from the bottom minimum point to the bisector of the line joining the two top points (Figure 2.6A). If the cycloid was interpreted as a chain of *i*'s, subjects bisected the "angle" of the cusp. More exactly, the midpoints of horizontal lines joining the opposite sides of the cusp were connected (Figure 2.6B). These midpoints form

Figure 2.6. Two different methods people use in estimating writing slant. If perceived as a *u*, subjects use method (A); if perceived as an *i*, method (B).

a straight lie with slant β given by:

$$\tan\beta = \frac{b}{a\cos\phi} \tag{2.5}$$

Surprisingly this slant is the same as computed in Figure 2.6A; this slant is also the same as the tangent at the cusp point.

This slant measure may be generalized to other writing shapes which do not have sharp top corners such as *e*'s. This measure corresponds to the slope of the velocity space ellipse at the vertical velocity zero crossing. The importance of this measure is that it shows writing slant is an artifact of an orthogonal oscillation system. The writing slant changes as the horizontal velocity a, the vertical velocity b, and the phase difference ϕ vary. Reexamining the difference in slant between tall and short letters in Figure 2.5, equation (2.5) yields an *e* slant of 61.8° and an *l* slant of 79.1°. The change in slant arises because the vertical velocity amplitude b changed.

Assuming that stylistic constraints require a constant writing slant, one must exercise care in going between tall and short letters to maintain slant. Since in letter height modulation the vertical velocity amplitude b is changed, it is required that the horizontal velocity amplitude a and phase difference ϕ also change to maintain a constant slant β. The constraints on letter shape and letter slant can be expressed by constants k_1 and k_2 respectively as follows.

$$k_1 = c - a\sin\phi$$
$$k_2 = \frac{b}{a\cos\phi} \tag{2.6}$$

From these two relations the letter height constraint implies that

$$b = k_2(c - k_1)\cot\phi \tag{2.7}$$

If a particular letter height is stipulated (namely $2b/\omega$), then the phase difference ϕ and the horizontal velocity amplitude a are determined. A plot of different e cycloids satisfying (2.7) but with different phases and velocity amplitudes is presented in Figure 2.7; Table 2.2 lists the relevant parameters. The perceived slant and shape seem to be constant among the several plots.

Table 2.2

Plot	a	ϕ	b
A	29.28	75	10.72
B	32.66	60	23.09
C	40.00	45	40.00
D	56.57	30	69.28
E	109.28	15	149.28

An alternating pattern of e's and l's that preserves both slant and shape under amplitude modulation alone is illustrated in Figure 2.8. Examining the velocity space ellipses, both the zero crossings and the slopes at the zero crossings are the same for the e and l ellipses.

The slant equation (2.5) is independent of frequency, and it would seem that if height control were obtained with frequency modulation alone the control of slant would be simplified. This idea works fine for the vertical oscillation, but the shape constraint forces the horizontal oscillation to modulate both frequency and acceleration amplitude. It will be seen in Chapter 4 that experimental measurements on humans show a combined amplitude and frequency modulation to control letter height, with frequency modulation the predominant influence.

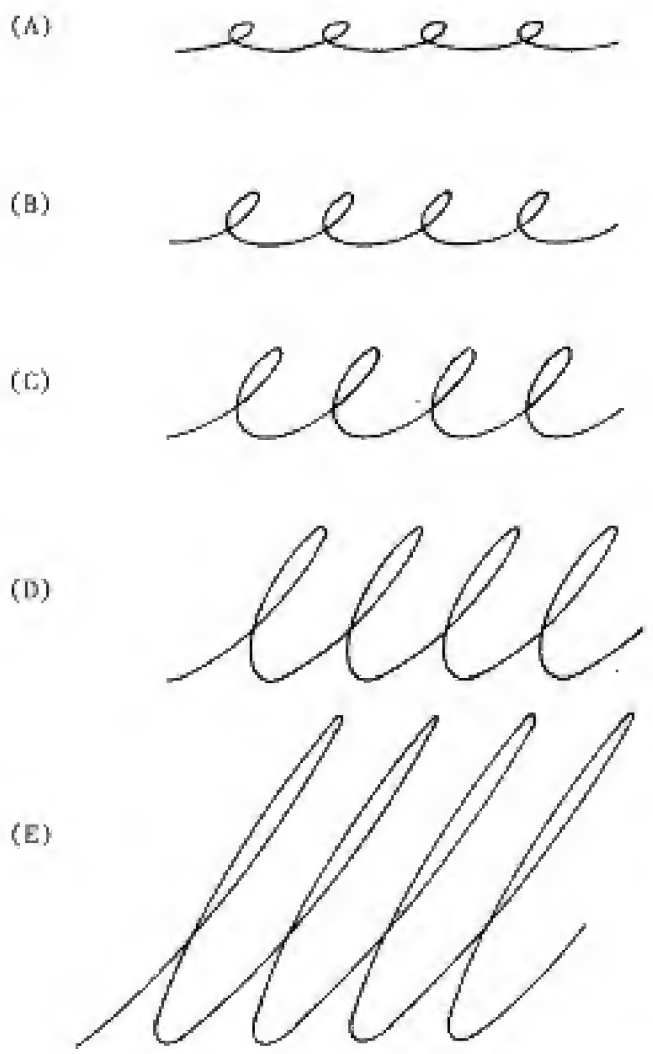


Figure 2.7. A sequence of e cycloids with constant slant measure β and constant vertical velocity zero crossing but different amplitude and phase values. The perceived slant and shape appear the same; the parameters for these cycloids appear in the text.

Handwriting Control: Manipulating the Velocity Zero Crossings

The previous discussion reveals that the vertical velocity zero crossing is of paramount importance in the control of handwriting (Figure 2.9). To control writing shape the value of the horizontal velocity at

Figure 2.8. Letter height modulation under constraints of constant slant and vertical velocity zero crossing is achieved with amplitude modulation. The velocity space ellipses on the left show that the vertical velocity zero crossing value and slope are the same for both *e* and *l* ellipses.

this zero crossing is altered. To control writing slant the slope of the velocity space diagram at the zero crossing is altered. Writing height is controlled by frequency modulation, which cannot be distinguished in the velocity space diagram, or by amplitude modulation, which affects the vertical elongation of the velocity space diagram.

2.3 Alternate Oscillation Patterns

The forcing function need not be sinusoidal for the main conclusions of this Chapter to hold. For example, a rectangular acceleration pattern (no rise time) yields triangular velocity profiles and a parallelogram in velocity space (Figure 2.10A). The writing produced by this oscillation pattern is a quite acceptable chain of *u*'s. Using once again the slope of the velocity space ellipse at the vertical velocity zero crossing as a measure of writing slant, the slant angle is given by $\tan\beta = b/a$. Psychophysical experiments for writing slant have not been performed with synthetic writing produced by rectangular acceleration patterns to ascertain the validity of this measure. If this measure is accurate, the independence of slant on phase is noteworthy. The advantage

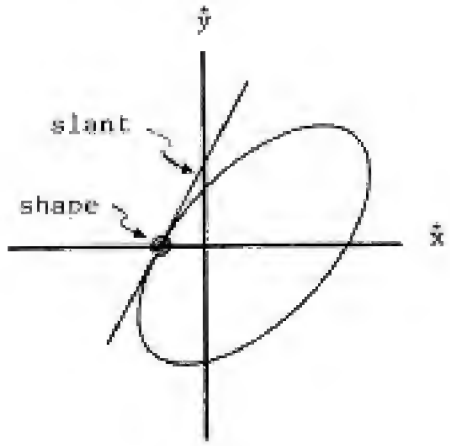


Figure 2.9. Manipulation of the velocity space diagram at the vertical velocity zero crossing leads to (1) control of shape by altering the value of the zero crossing, and (2) control of slant by altering the slope at the zero crossing.

of phase independent writing slant is a considerable simplification of control.

Trapezoidal peaks have also been proposed as models of the acceleration patterns in handwriting [McDonald 1966]. The slopes to the plateaus presumably model a linear rise time of the actuation. The effect of the linear rise time on the velocity space diagram is to round the corners of the rectangular pattern's velocity space parallelogram (Figure 2.10B). If the rounding does not occur at the vertical velocity zero crossing, the writing slant is again given by $\tan\beta = b/a$. If the rounding occurs through the zero crossing the writing slant becomes phase dependent as for the sinusoidal case.

2.4 Nonorthogonal writing axes

Other explanations for writing slant have been put forth. Mermelstein [1964] proposed that a nonorthogonal axes disposition could explain writing slant. The slant of the handwriting would follow the direction of the slanted vertical axis. From the previous discussion it is clear that coupled oscillations cause as a side effect slanted writing, and that therefore a nonorthogonal joint disposition cannot alone account for writing

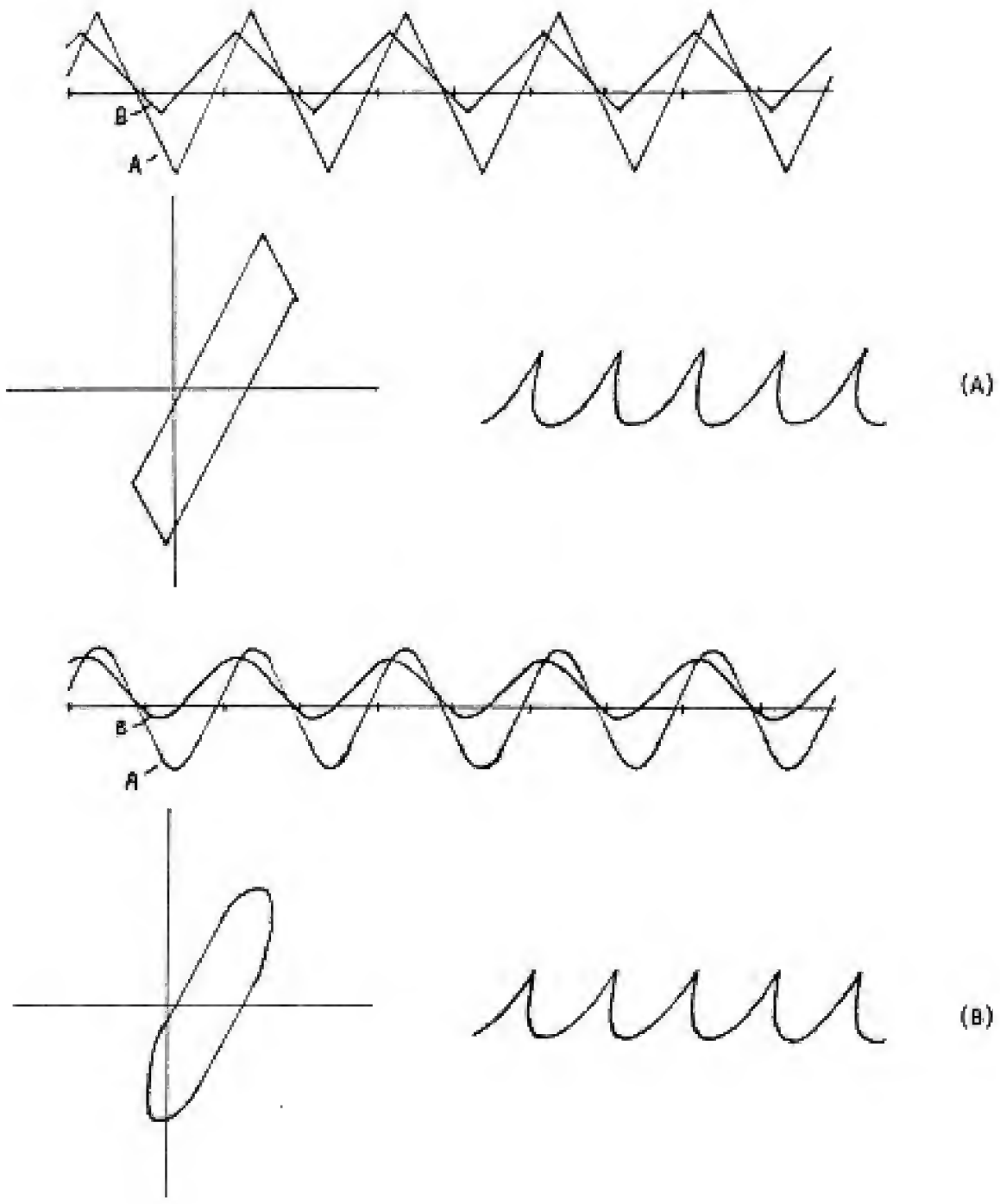


Figure 2.18. (A) Triangular velocity profiles resulting from a rectangular acceleration pattern; (B) the velocity profiles are rounded with a trapezoidal acceleration pattern.

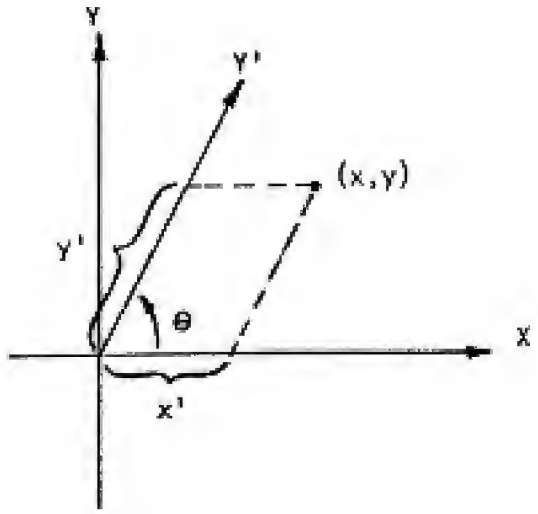


Figure 2.11. Shear transformation for a nonorthogonal joint angle θ. The equations governing this transformation are: $x' = x - y\cot\theta$. $y' = y\csc\theta$.

slant. Any slant from nonorthogonal axes must be added to the slant caused by coupled oscillations.

There is a question as to how one can distinguish these two contributions. If θ is the angle the vertical axis makes with the horizontal axis, and if cycloidal velocity equations (2.1) act along the nonorthogonal axes, then the orthogonally measured velocities are (see Figure 2.11):

$$\begin{aligned} \dot{x} &= a\sin(\omega t + \phi) + c + b\cos\theta\sin\omega t \\ \dot{y} &= b\sin\theta\sin\omega t \end{aligned} \tag{2.8}$$

The velocity space diagram corresponding to (2.8) is also an ellipse but with altered tilt and eccentricity. It is impossible to factor the product $b\cos\theta$ by manipulations on the measured velocities. That is to say, there are infinitely many pairs (b, θ) that yield exactly the same writing. This multiplicity may actually pose an advantage for situations where the joint disposition changes within a word; for example, in going between extremes of a joint before repositioning. The writing however can be kept the same merely by adjusting b. The issue of nonorthogonal axis disposition is discussed further in the next chapter. For the present the extent of

any contribution to slant from nonorthogonal axes remains an open question.

Another explanation for writing slant was put forth by Pellionisz and Llinas [1979], who proposed that writing slant arose from computing delays in the cerebellum. According to their theory an idealized up-down movement gets transformed into a slanted line because the trajectory computation lags the actual trajectory.

Chapter 3. Handwriting under a Spring Model

The oscillation theory in the last chapter indicates how handwriting can be produced. The next step is to propose how the theory could be implemented specifically by the human motor system. Particular issues that have to be resolved are the agent responsible for the oscillation, the nature of the control variables, and the influence the control variables have on bringing about the various modulations.

A recent theory of motor control likens muscle to a spring system with variable zero setting [Feldman 1974a, 1974b]. This spring muscle model has a particular affinity to the oscillation theory of the last chapter, and provides a useful framework for considering one way in which the human motor system might implement this theory. In this chapter the issues of parameter control raised in the previous paragraph are developed in the context of the spring muscle model.

3.1 The Spring Muscle Model

The spring muscle model is a simplification of the length-tension curves of muscle. In Figure 3.1 length-tension curves under isometric contraction at several firing rates for the cat soleus muscle are shown [Rack and Westbury 1969]. There is some question as to what portions of the length-tension curves are used in actual movement. Some authors [Zierler 1974, Hill 1970, Cook and Stark 1967] maintain that the active portions occur near the length at which there is maximum isometric tension and which Zierler also calls the rest or natural length of muscle. Cullins et al. [1975] on the other hand report that the linear portions at short muscle lengths are used in eye movement; Feldman [1974a, 1974b] in his model of movement also assumes the linear portions are active.

The spring muscle model is derived by assuming muscle operates at the linear short length regions. The length-tension curves in this region are also assumed parallel, although Figure 3.1 shows a slightly increasing slope with firing rate. The simplified length-tension curves of agonist and antagonist muscles would intersect and overlap as in Figure 3.2; different length-tension curves are selected by adjusting the muscle firing rate.

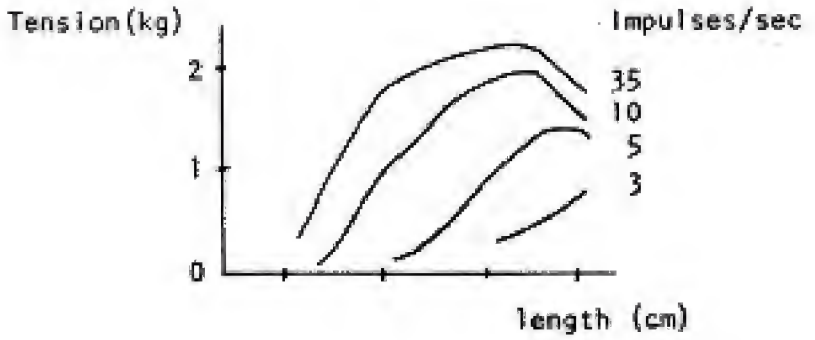


Figure 3.1. Length-tension curves from the cat soleus muscle, from Rack and Westbury 1969.

Feldman [1974a, 1974b] has proposed that movements are executed by selecting a pair of agonist-antagonist length-tension curves that intersect at the desired position. This process is illustrated in Figure 3.2. Suppose the system is currently at length L_0 under innervation rates g_0 for the agonist and n_0 for the antagonist. If the innervation rate of the agonist is changed to g_1, a different agonist length-tension curve is selected and the equilibrium length shifts to L_1. Assuming no delay in tension development and ignoring velocity effects, the arrow in the Figure indicates the tension course. There is an isometric buildup of tension from P_0 to P_2 followed by a decay to P_1, where the tension in agonist balances the tension in antagonist.

The curves in Figure 3.2 lead to a model of muscle as a spring with variable zero setting. The slope K of the curves represents the spring constant, and the variable zero setting L_z corresponds to the selection of firing rate. The force exerted by a muscle is thus $K(L - L_z)$. This simplified muscle model is schematized in Figure 3.3, which includes a passive damping factor b. The equation of motion for the spring system of Figure 3.3 is:

$$m\ddot{x} = -b\dot{x} + k_p(x_p - x) - k_n(x - x_n) \tag{3.1}$$

In the remaining portions of this chapter we neglect viscous and passive elastic components of muscle in order to develop the main points.

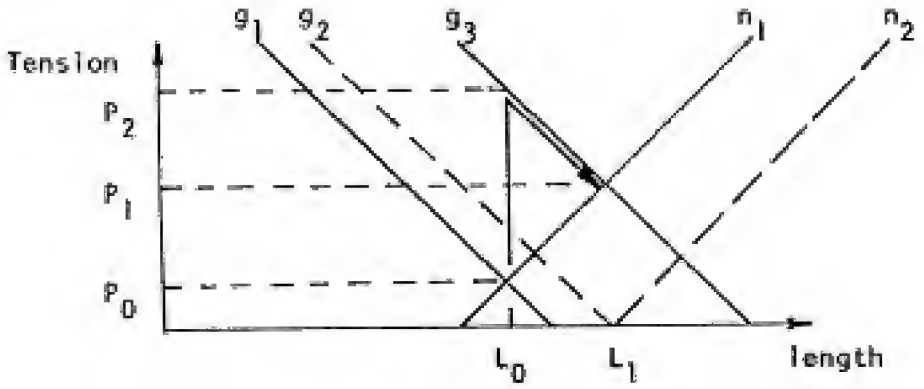


Figure 3.2. The equilibrium point of the intersecting length-tension curves of agonist (g labels) and antagonist (n labels) shifts from L_0 to L_1 when the firing rate of the agonist is raised from g_1 to g_3 and the antagonist rate remains at n_1.

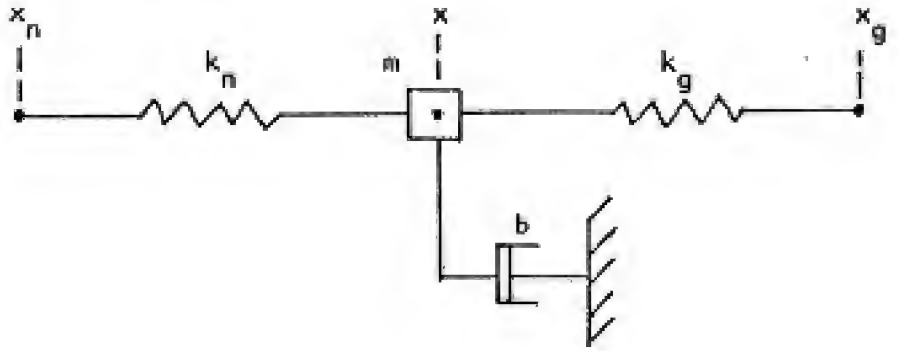


Figure 3.3. A simplified muscle model with the following parameters: b is the coefficient of passive damping, k_g is the spring constant of the agonist, k_n is the spring constant of the antagonist, m is the mass, x is the mass position, x_g is the agonist variable zero setting, x_n is the antagonist variable zero setting.

Handwriting with Orthogonal Spring Pairs

If handwriting is executed by two orthogonal joints, we can model the system by the action of two

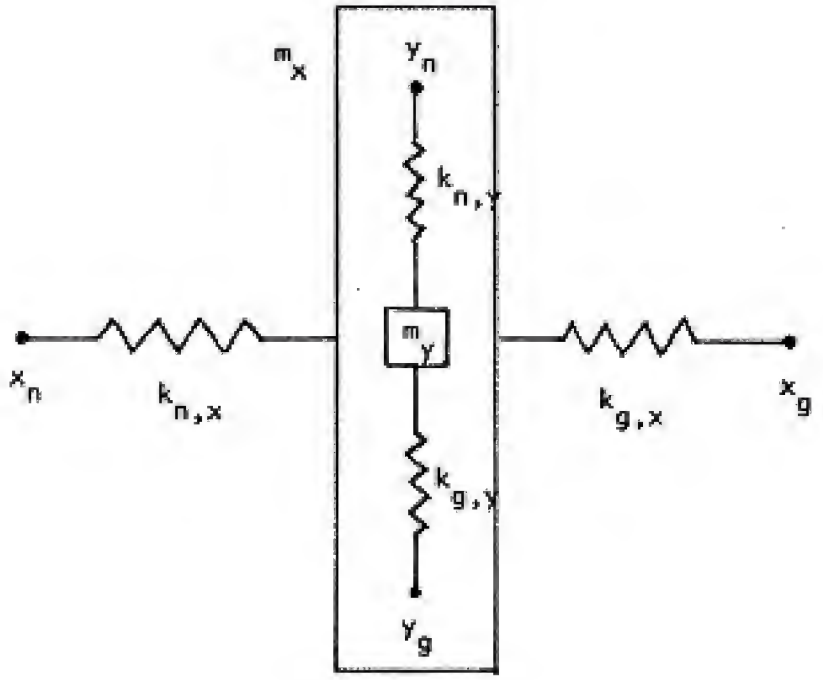


Figure 3.4. Two orthogonal pairs of springs serve as a model for handwriting. The y springs act on the pen mass m_y, the x springs act on the mass m_x which is the sum of the pen mass and the y spring platform.

orthogonal opposing pairs of springs (Figure 3.4). The y springs act on the pen mass m_y; the x springs act on the pen mass and on the y spring masses themselves, indicated by attachments to a platform of total mass m_x mounted with the pen mass and the y springs. Neglecting viscous friction, the equation of motion of the pen mass in the y direction is:

$$m_y\ddot{y} = k_{g,y}(y_g - y) - k_{n,y}(y - y_n). \tag{3.2}$$

Absorbing the mass m_y into the spring constants and solving this equation by Laplace transforms:

$$y(t) = \frac{\dot{y}(t_0)}{\omega_y} \sin \omega_y(t - t_0) + (y(t_0) - \frac{k_{g,y}y_g + k_{n,y}y_n}{\omega_y^2}) \cos \omega_y(t - t_0) + \frac{k_{n,y}y_g + k_{n,y}y_n}{\omega_y^2}, \tag{3.3}$$

where $\omega_y^2 = k_{g,y} + k_{n,y}$. It is also more convenient to rewrite (3.3) in the following form:

$$y(t) = \frac{-b}{\omega_y} \cos(\omega_y(t - t_0) + \phi_y) + \frac{k_{g,y}y_g + k_{n,y}y_n}{\omega_y^2}, \tag{3.4}$$

where

$$b \sin \phi_y = \dot{y}(t_0),$$
$$b \cos \phi_y = \omega_y(\frac{k_{g,y}y_g + k_{n,y}y_n}{\omega_y^2} - y(t_0)).$$

According to the theory of the last chapter the horizontal movement has a constant velocity sweep superimposed on an oscillation. The justification as discussed in the next chapter comes from the observation that the velocity space diagrams stay centered about the same point through all the different modulations of velocity for different letter shapes. The constant velocity sweep is imposed on any additional change to the x spring zero settings. Setting $x_g(t) = x_g + c(t - t_0)$ and $x_n(t) = x_n + c(t - t_0)$ and solving the equations corresponding to (3.2),

$$x(t) = \frac{-a}{\omega_x} \cos(\omega_x(t - t_0) + \phi_x) + \frac{k_{g,x}x_g + k_{n,x}x_n}{\omega_x^2} + c(t - t_0). \tag{3.5}$$

where $\omega_x^2 = k_{g,x} + k_{n,x}$ and

$$a \sin \phi_x = \dot{x}(t_0) - c,$$
$$a \cos \phi_x = \omega_x(\frac{k_{g,x}x_g + k_{n,x}x_n}{\omega_x^2} - x(t_0)),$$

A Mechanical Agent for the Oscillation

When differentiated, equations (3.4) and (3.5) are the same as the velocity sinusoids of equations (2.1) and (2.2). A sinusoidal oscillation arises from the spring muscle model as the simplest method of operation. Once

an appropriate set of initial conditions is set up, the oscillation propagates indefinitely with the zero settings unchanged. With the orthogonal spring pairs, a cycloid ensues in the manner of Figure 2.2. The agent for the oscillation under the spring model is a purely mechanical one, assuming of course no dissipation from viscous elements. If viscous elements were to be included, an active involvement of the control centers to maintain the mechanical oscillation would be required.

3.2 Modulating Letter Shape and Height with Springs

The previous section indicated that controlling the vertical velocity zero crossings and controlling the vertical velocity amplitudes of a cycloid result in modulation of shape and modulation of height. The vertical velocity zero crossing was shown to be a function of the horizontal velocity amplitude and phase. In the spring model the amplitudes and phases are controlled by adjusting the zero settings of the springs. A difficulty under this spring model is that phase and amplitude cannot be controlled independently. In order to achieve a particular zero crossing, for example, one has to search for the particular zero setting applied at a particular time that yields the desired product $a \sin \phi$. In order to modulate letter height the modulation is best applied at the bottom corner at zero vertical velocity; the resulting vertical amplitude changes without phase change.

It is a potential difficulty with the spring model if a physical situation exists with $\omega_x \neq \omega_y$ and there is no way to make these frequencies equal. For example, the mass of the writing implement, the size of the limbs involved, and the frictional contact with paper are subject to change and influence the frequencies. This difficulty implies a need to be able to adjust frequencies and hence to vary the spring constants. A variable spring constant model is considered in section 3.3. For the present discussion we assume fixed spring constants with $k = k_{g,x} = k_{a,x} = k_{g,y} = k_{a,y}$ and $\omega = \omega_x = \omega_y$.

A final issue before considering corner shape and letter height modulation is that a change in zero setting can be applied to either the agonist or antagonist spring or to both at the same time. As far as the mathematics is concerned these situations are all equivalent. Examining for example the y equation (3.4), the influence of the zero settings is given by $(y_g + y_a)/2$ after taking into account the simplifications of the previous paragraph.

Thus for a given axis there is only one functional control parameter, corresponding to the combined change in zero settings of both springs. There do arise situations where a zero setting change can be applied only to one spring in order to avoid the other spring from pushing as well as pulling. Such situations will not arise in the following examples, and for simplicity we will assume the change in zero setting Δy or Δx is applied to the agonist spring.

Modulation of corner shape

To illustrate that the vertical velocity zero crossing can be manipulated by changing the horizontal agonist zero setting, the increments Δx to x_g required to produce the *ecme* of Figure 2.4 are presented. The *e* cycloid is generated with initial conditions in Table 3.1. The three increments Δx that bring about phase and amplitude changes in Table 2.1 corresponding to the *u*, the *n*, and the *e* are 1.967, 2.0, and 3.493 mm respectively. In fact the writing in Figure 2.4 was generated by the spring model.

Table 3.1			
Parameter	Value	Parameter	Value
$x(t_0)$	-1.5 mm	$y(t_0)$	-2.8 mm
$\dot{x}(t_0)$	65.0 mm/sec	$\dot{y}(t_0)$	0.0 mm/sec
x_g	4.0 mm	y_g	5.0 mm
x_n	-4.0 mm	y_n	-5.0 mm
c	20.0 mm/sec	ω	5.0 cycles/sec

Modulation of Letter Height

We consider here how to produce the amplitude modulation for letter height that yields the alternating *e*'s and *l*'s of Figure 2.8. There are three constraints on the letter height modulation: (1) the achievement of a particular letter height, in this case a ratio of $2\sqrt{2}$ (close to the human average), (2) a constant writing slant, and (3) a constant corner shape. Fortunately there are an equal number of control variables, namely the zero setting changes Δx and Δy, and the time t_1 of the Δx application. The interletter separation is fixed by these choices of parameters and is not under independent control.

As mentioned in Chapter 2 the y modulations are best applied at points of zero vertical velocity, because the phase change is then always zero and the maintenance of an even baseline is simplified. Under acceleration amplitude modulation for letter height the writing size is directly proportional to the acceleration amplitude. The time of horizontal modulation is found to coincide with the point of vertical modulation. The increments Δy and Δx to the vertical and horizontal agonist springs to produce the first l of Figure 2.8 is presented in Table 3.2, along with the amplitudes and phases of the corresponding cycloids.

Table 3.2

Parameter	$t = 0.0$ sec	$t = 0.2$ sec
Δy	0.0 mm	10.24 mm
Δx	0.0 mm	5.37 mm
b	28π mm/sec	248.8 mm/sec
a	65.16 mm/sec	140.7 mm/sec
ϕ	0.762 rad.	0.326 rad.

The slant of the resultant l is the same as the slant of the e and the vertical velocity zero crossing has the same value for the l and the e.

Combined Height and Corner Shape Modulation

Ordinarily when modulating for height in upper zone letters such as l or b no special corner shaping need be done by the horizontal springs. For lower zone letters such as y and j on the other hand a considerable degree of corner shaping is required. Figure 3.5 shows that just a height modulation for a lower zone letter does not produce an acceptable shape, unlike a height modulation for an upper zone letter which produces an acceptable l.

Obtaining a lower zone loop as in the letter y (Figure 3.6) involves a transition from counterclockwise to clockwise movement, then back from clockwise to counterclockwise movement. A difficulty resulting from these transitions is creating enough horizontal separation between the y and the next letter e, and so the bottom of the y must curl around to make the upstroke more propitious. To achieve these transitions and spacing requires 4 horizontal zero setting increments Δx in a span of 150 msecs, bringing about large changes

Figure 3.5. Letter height modulation produces both upper and lower zone letters, but the lower zone letters require x modulation for shaping.

in phase and very large horizontal velocity amplitudes (Table 3.3).

Table 3.3			
Time	Δx	$\Delta\phi$	a'
0.43 secs	2.0 mm	24.5 degrees	67.5 mm/sec
0.74 secs	-11.0 mm	67.0 degrees	164.6 mm/sec
0.80 secs	11.0 mm	-24.6 degrees	309.6 mm/sec
0.85 secs	-12.0 mm	-47.7 degrees	312.7 mm/sec
0.89 secs	11.0 mm	-21.7 degrees	74.5 mm/sec

Transitions between Counterclockwise and Clockwise Movement

If one can speak about the relative difficulty in producing various letters, the number of transitions between counterclockwise and clockwise movements would indicate the level of difficulty. Associated with such a transition are usually large phase and amplitude changes, and one might construct a complexity measure based on these changes. By any such measure lower zone loops as in y would have to be considered difficult. Another difficult form is the letter h (Figure 3.7), which requires two counterclockwise/clockwise transitions.

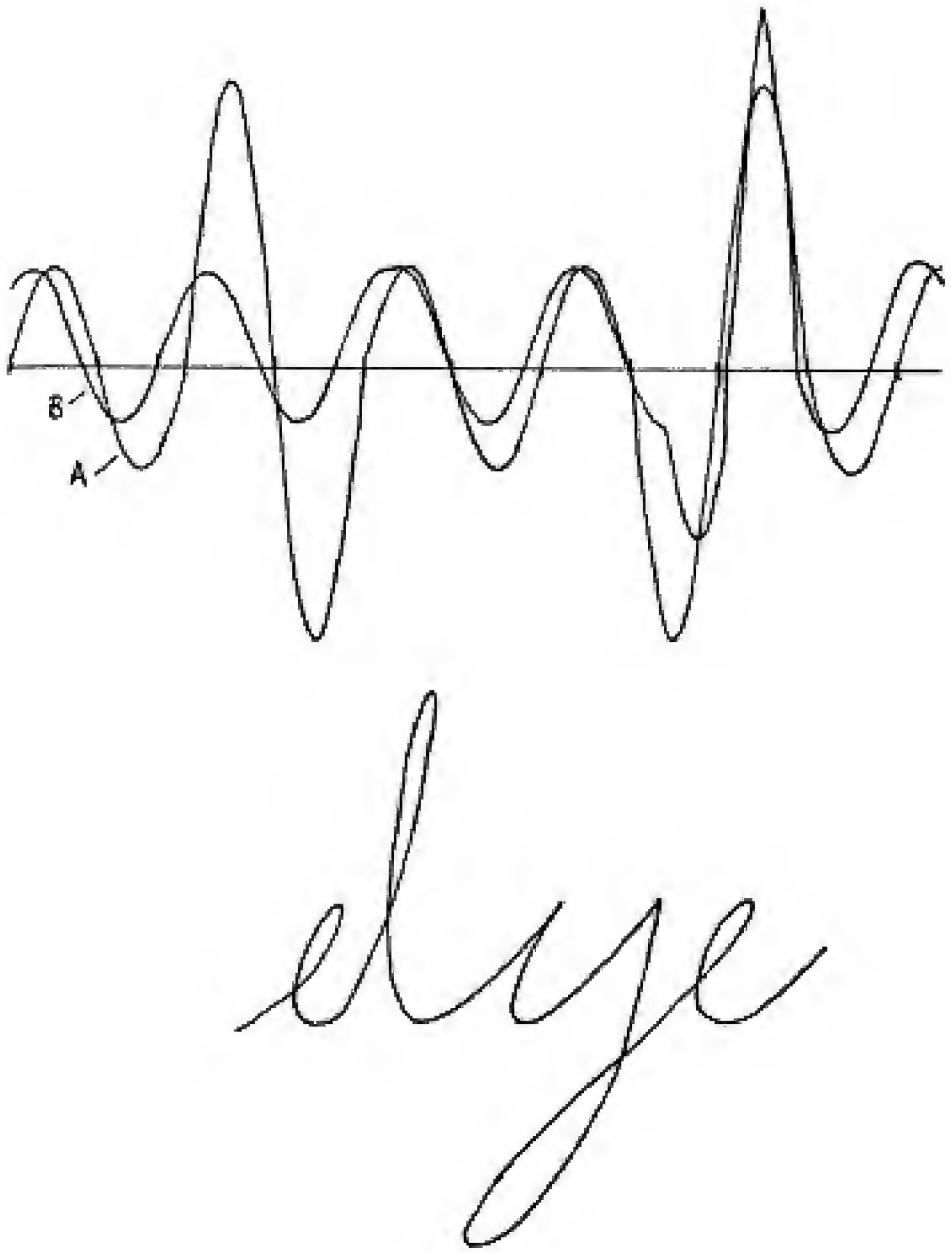


Figure 3.6. Amplitude modulation is used to produce short and tall letters in this sequence **edge**. Note the change in slant between short and tall letters.

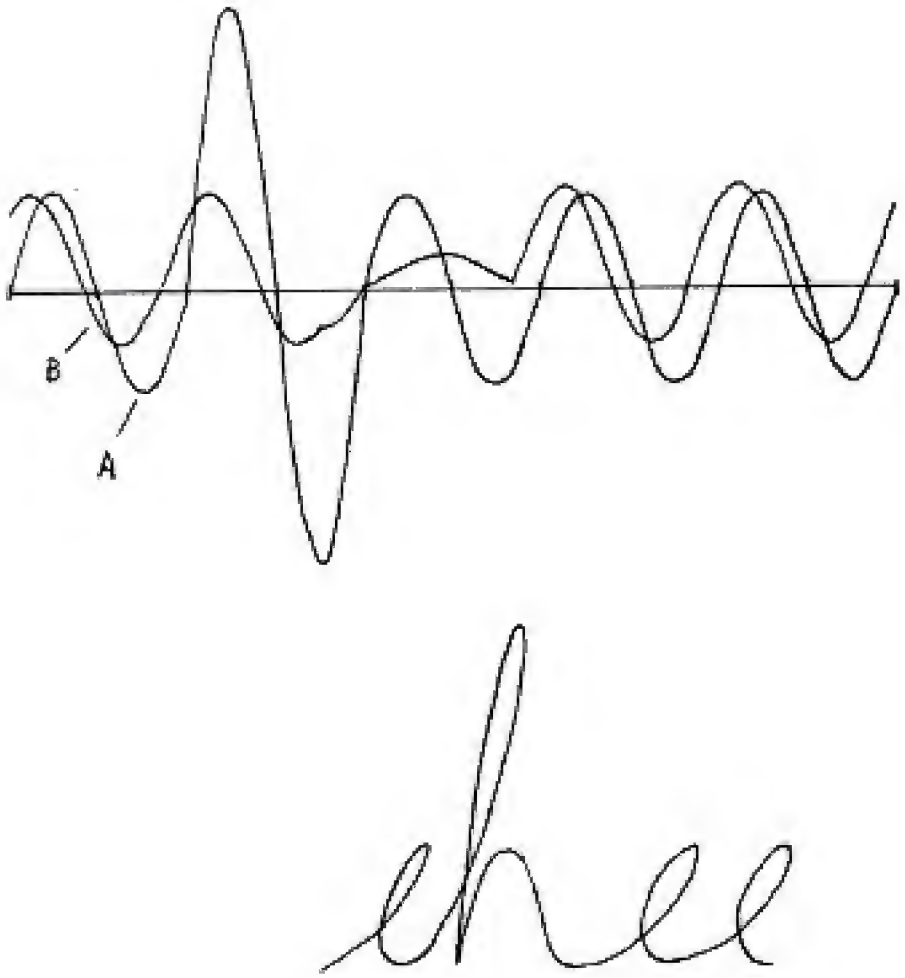


Figure 3.7. The letter *h* presents problems in production because of the need for several large phase changes.

The first transition goes into making the bottom cusp coming from the top loop, accomplished with 2 closely spaced modulations resulting in fairly large phase changes (table 3.4). The second transition restores the sequence of e's, but at the cost of a very large phase change.

Table 3.4			
Time	Δx	$\Delta\phi$	a'
0.35 secs	-5.0 mm	65.2 degrees	50.7 mm/sec
0.40 secs	-2.0 mm	55.2 degrees	20.3 mm/sec
0.57 secs	5.0 mm	-113.0 degrees	69.3 mm/sec

In Chapter 4 some evidence will be presented that fast human handwriters achieve speed by simplifying the script to eliminate all clockwise movements. If one may draw an analogy between spring writing and human writing, the reason might be that counterclockwise/clockwise transitions are too difficult to make fast.

3.3 Variable Spring Constant Model

The previous discussions in this chapter took place in the context of a spring muscle model with fixed stiffness. There is reason to presume that the human motor system can control not only muscle tension but also muscle stiffness. There are several reasons why stiffness regulation would be desirable from the standpoint of the spring handwriting model. (1) There is a need to equate horizontal and vertical oscillation frequencies under different physical conditions, as mentioned earlier. (2) Stiffness regulation would provide another control variable and would allow more flexibility in modulation. (3) If letter height could be modulated by frequency instead of by amplitude, then the slant equation (2.5) implies that the slant would not change because slant is not a function of frequency. Chapter 4 provides evidence that people use in part a strategy of frequency modulation for letter height.

We take as an analytic model of stiffness variation with slant an extrapolation from the length-tension curves of Rack and Westbury [1969]. If the linear portions of these curves are extended they seem to meet at a point (Figure 3.8). For the horizontal antagonist spring, for example, the dependence of the spring constant $k_{n,x}$ of a variable stiffness spring corresponding to this ray of straight lines is:

$$k_{x,n} = \frac{T_{n,x}}{x_n - x_{n,0}} \tag{3.6}$$

where $x_{n,0}$ is the length at which the rays intersect and $-T_{n,x}$ is the tension at $x_{n,0}$. It should be emphasized

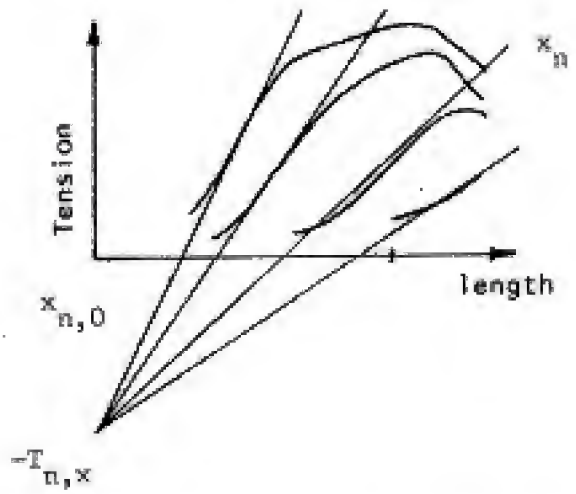


Figure 3.8. The length-tension curves of Rack and Westbury (1969) when extended meet at a point. The dependence of the slope $k_{n,x}$ on on the adjustable zero setting z_n can be characterized by $T_{n,x}/(z_n - z_{n,0})$.

that the exact dependence of stiffness on zero setting is not important for analyzing handwriting under a variable spring constant model, and the relation (3.6) is chosen merely because it is convenient and because the literature does not contain an alternative expression.

We reexamine the issue of control of letter height, slant, and shape raised in section 3.2. A set of initial conditions appropriate for this model are presented in Table 3.5.

Table 3.5

Parameter	Value	Parameter	Value
$x_{g,0}$	20.0 mm	$y_{g,0}$	20.0 mm
$x_{n,0}$	-20.0 mm	$y_{n,0}$	-20.0 mm
x_g	17.5 mm	y_g	17.5 mm
x_n	-17.5 mm	y_n	-17.5 mm
$T_{g,x}$	1233.7 mm/sec^2	$T_{g,y}$	1233.7 mm/sec^2
$T_{n,x}$	1233.7 mm/sec^2	$T_{n,y}$	1233.7 mm/sec^2
$x(t_0)$	-0.51 mm	$y(t_0)$	-1.0 mm
$\dot{x}(t_0)$	23.2 mm/sec	$\dot{y}(t_0)$	0.0 mm/sec
c	7.15 mm/sec		

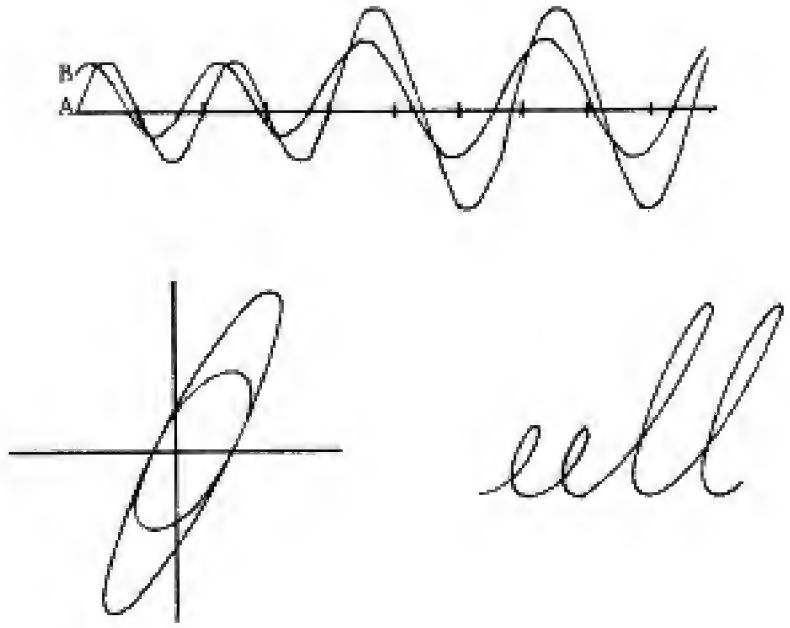


Figure 3.9. Letter height control by combined frequency and amplitude modulation in a variable spring model is illustrated by the synthetic writing in the bottom right diagram. Top diagram: vertical (A) and horizontal (B) velocity traces, showing frequency and amplitude modulation. Bottom left: velocity space diagram shows slant and shape constraints have been met.

With these parameters a train of *e*'s is produced with the same shape as produced by the fixed spring model and the parameters of Table 3.1. When modulating for letter height human subjects typically decrease the frequency by a factor of $\sqrt{2}$ and double the vertical velocity amplitude (see Chapter 4 for the experimental data). The increments Δy_g and Δy_a to the agonist and antagonist y springs can be chosen to satisfy both conditions (Table 3.6). When modulating the horizontal springs to preserve slant and shape, the horizontal frequency must also be made to match the vertical frequency. The horizontal modulation is forced by the mathematics to occur at the same time as the vertical modulation, and the increments Δx_g and Δx_a to the agonist and antagonist x springs are given in Table 3.6. The resultant writing appears in Figure 3.9. The shape and slant constraints have been met. Table 3.6 shows that the x and y frequencies have been decreased by a factor of $\sqrt{2}$, and the vertical velocity amplitude has been doubled.

Table 3.6		
parameter	0.0 sec	0.4 sec
Δy_{0}	0.0 mm	-2.08 mm
Δy_{t0}	0.0 mm	3.00 mm
$k_{0,y}$	493.5 /sec	269.3 /sec
$k_{t0,y}$	493.5 /sec	224.2 /sec
ω_y	10π /sec	$5\pi\sqrt{2}$ /sec
b	10π mm/sec	20π mm/sec
ϕ	$\pi/4$ rad.	0.464 rad.
Δx_{0}	0.0 mm	-2.28 mm
Δx_{t0}	0.0 mm	2.75 mm
$k_{0,x}$	493.5 /sec	258.2 /sec
$k_{t0,x}$	493.5 /sec	235.2 /sec
ω_x	10π /sec	$5\pi\sqrt{2}$ /sec
a	22.7 mm/sec	35.9 mm/sec

A frequency decrease by $\sqrt{2}$ and a doubled vertical velocity amplitude implies that the vertical velocity amplitude increase is due equally to a frequency modulation and an acceleration amplitude modulation. That is to say, a constant acceleration amplitude is not maintained through a height modulation. Nevertheless the idealization of achieving height modulation without slant change by frequency modulation would have been obtained if the horizontal velocity amplitude could have been doubled also and the phase difference kept constant. Unfortunately the shape constraint prevents this easy solution for slant control, since equation (2.6) requires the phase difference to change and the horizontal velocity amplitude to assume some value other than double the original amplitude.

Nevertheless there has been a net benefit in reducing the "difficulty" of letter height modulation because both the horizontal and vertical velocity amplitude changes are half that in the fixed spring model, and the required phase change is smaller.

3.4 Coactivation Versus Alternate Activation

So far the spring model has been presumed to work by having both springs exert force at the same time. It is conceivable to have only one spring on at a time. In the simplest mode of operation the zero settings x_0

and x_0 or y_0 and y_0 would be equal, and whenever the position passes through the zero setting the agonist spring is switched off and the antagonist spring on. The equations of motion are exactly the same as before, except that the frequency is smaller by a factor of $\sqrt{2}$. The control of both systems therefore is equivalent. It is possible to have a mode of operation where the antagonist spring comes on before the point of maximum velocity, but analysis of such controls shows that maintenance of an even baseline is rather difficult because of the complexity of the equations.

It is possible to use a scheme of switching between simultaneous and alternate agonist/antagonist spring activation to achieve a limited frequency modulation under a fixed spring constant model. A frequency decrease by $\sqrt{2}$ would by itself result in a doubled letter height. Any additional height modulation would require changing the vertical acceleration amplitude.

Chapter 4. Experimental Measurements with Humans

In this chapter an apparatus designed for accurate measurement of position, velocity, and acceleration during human handwriting is described. Measurements of human handwriting with this apparatus are then compared against the model of the previous two chapters.

4.1 The Experimental Apparatus

Acceleration Measurement

Measurements of acceleration during handwriting are obtained with accelerometers mounted on a 6 degree of freedom X-Y sliding rail system (Figure 4.1). A linear bearing housing slides along the X-axis rod attached to a metal base. At the same time the Y-axis rods may slide through the housing. There are three degrees of freedom at the pen holder: two stacked axial bearings for Y-axis rotation, a hinge joint for rotation with respect to the end of the Y-axis rods, and a bronze bushing for rotation of the pen inside the holder. Play in the apparatus is negligible.

The whole apparatus allows nearly frictionless movement. Subjects reported no constraints on their writing movement and no difficulties in grasping the pen near the holder. Tests showed that the inertia of the apparatus did not affect the writing act. The finger and wrist muscles are evidently overpowered for the handwriting movement.

The acceleration is sampled at a rate of 250 coordinate pairs per second at a resolution of better than 0.01 g's. A simple RC filter to the accelerometers with cutoff frequency 25 Hz removes high frequency paper noise. It is followed by a triangular digital filter with Nyquist frequency 25 Hz.

This apparatus has some advantage over previous acceleration measurement apparatuses. Designs with accelerometers or force sensors mounted in the pen [Crane and Savoie 1977, Herbst and Liu 1977] are susceptible to orientation problems. People roll the pen in their fingers as they write and also change the pen

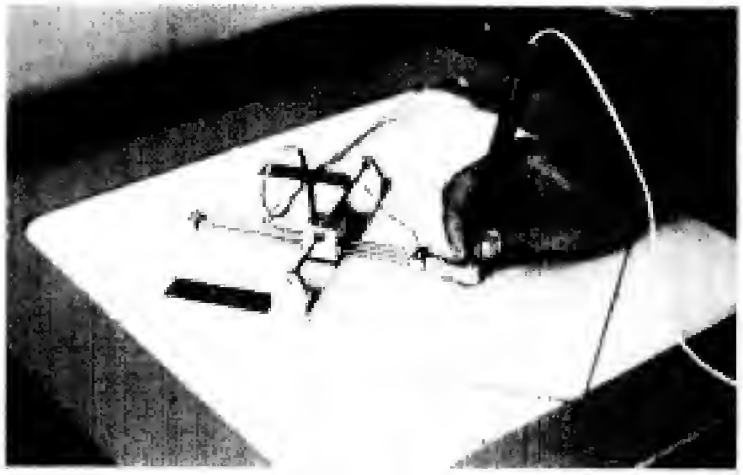


Figure 4.1. A six degree of freedom x-y rail system for measurement of acceleration and position.

inclination; the sensitive axes of the accelerometers or force sensors change unpredictably and the data loses its meaning. Constraints imposed on subjects to obviate orientation problems interfere with the writing act. A second class of apparatuses derive acceleration from position; these include writing tablets, electrolytic water tanks [McDonald, Yasuhara], teledelthos spark systems [Denier van der Gon and Thuring], and a sliding rail system with potentiometers [Koster and Vredenbregt]. The time and spatial resolution of these apparatuses are not currently sufficient for accurate acceleration derivation; there is no choice but to measure acceleration directly. The sliding rail system employed in this thesis obviates the orientation problems with pen mounted systems, and provides an accurate and detailed record of the acceleration of the pen tip.

Position Measurements

Position measurements are obtained with a Summagraphics ID Data Tablet/Digitizer. The writing tablet pen fits into the holder of the accelerometer apparatus. Velocity is estimated by differentiating position.

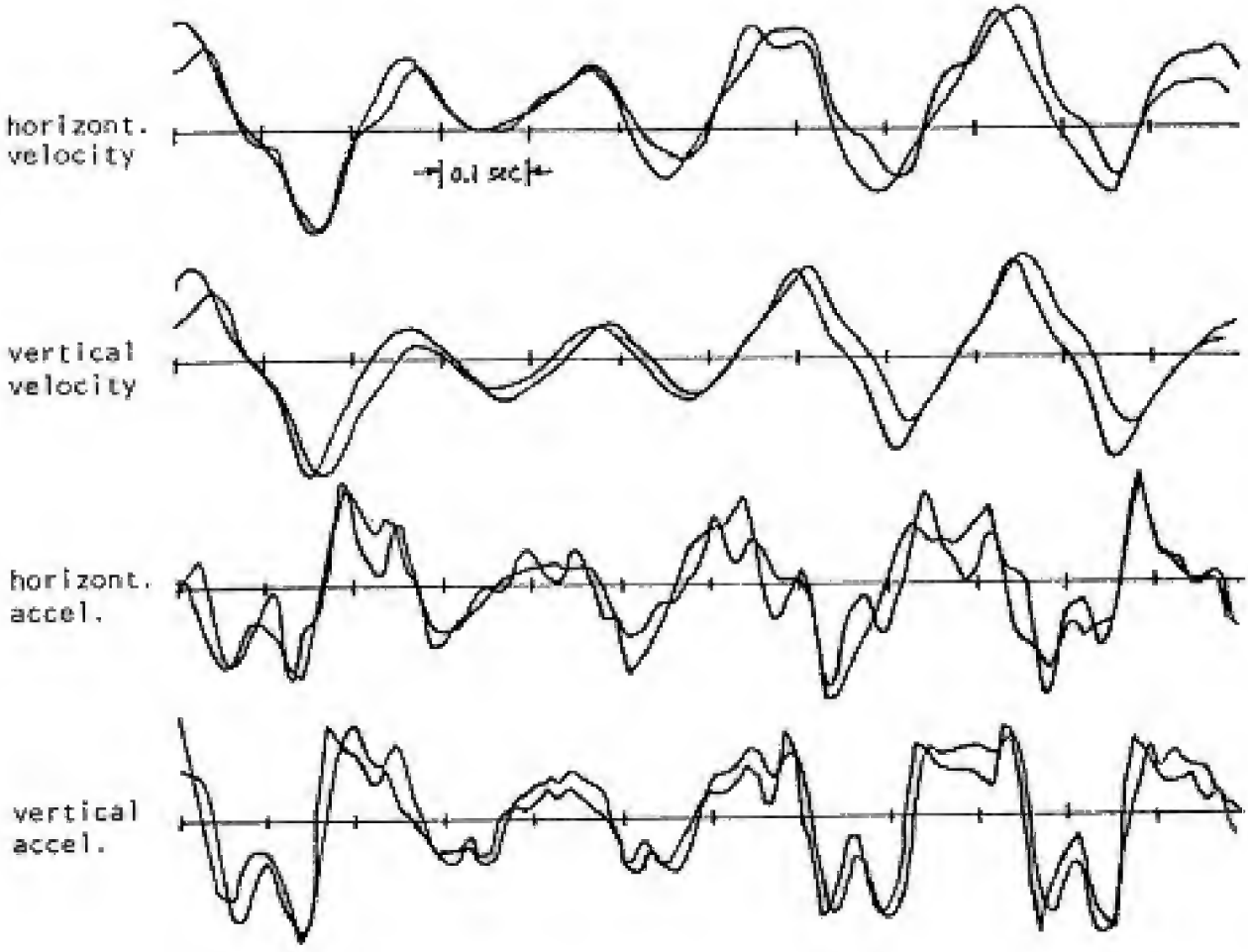


Figure 4.2. Velocity curves derived from position and from acceleration show good agreement. Doubly differentiated position agrees well with measured accelerations.

Velocity estimated by integrating acceleration shows close agreement to the position derived velocities (Figure 4.2). Also present in the Figure is the matching of measured acceleration to doubly differentiated position, again showing a good agreement. The sampling rate for position is 94 coordinate pairs per second at a spatial resolution of 0.1 mm. A narrow triangular filter was applied to the position data to yield a smoothed velocity curve. A complete recording of a handwriting movement is illustrated in Figure 4.3.

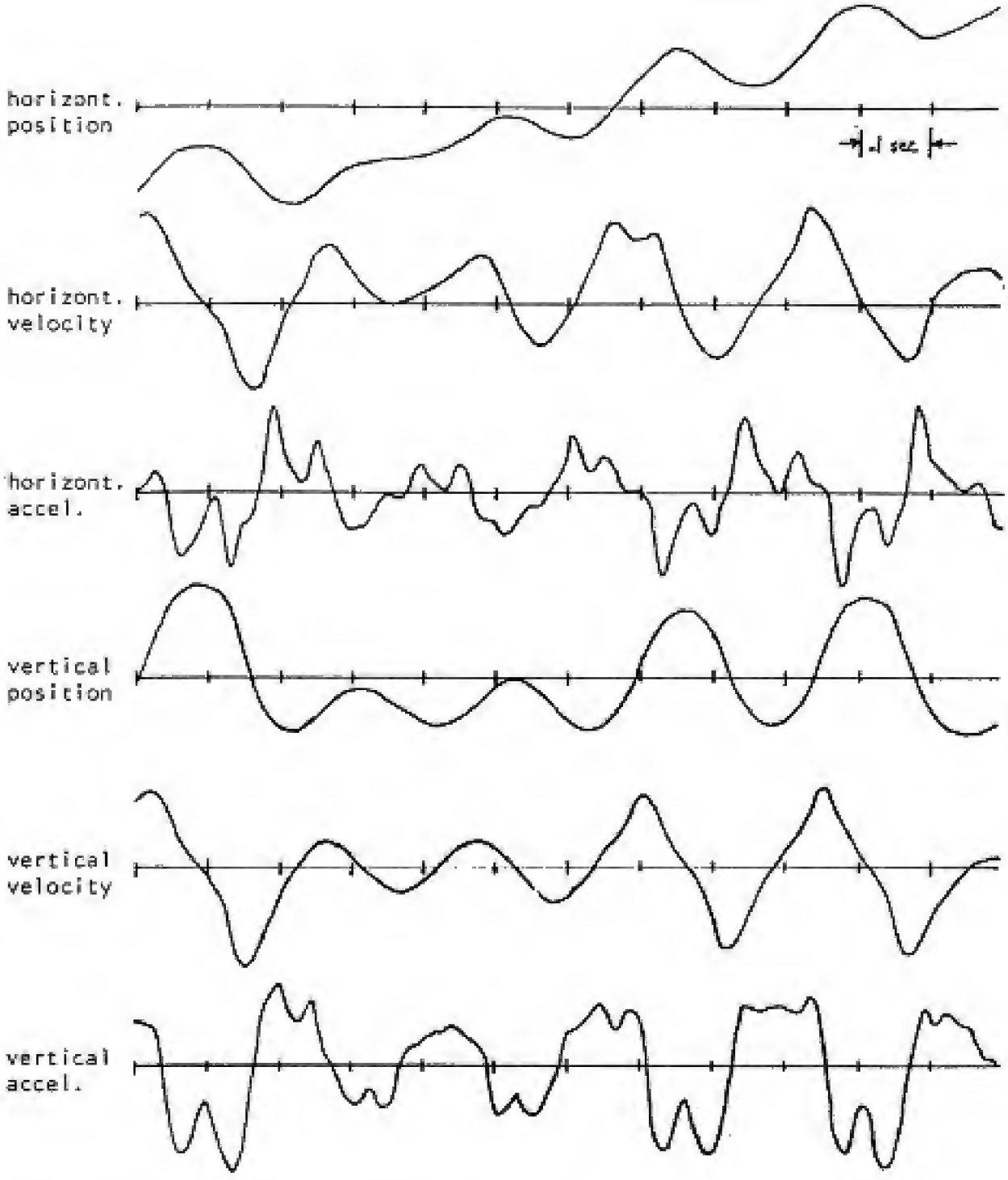


Figure 4.3. A complete recording of *hell* by subject KAS.

Figure 4.4. Acceleration profile for a down-up finger movement on plexiglass by subject HOL.

The sharp negative acceleration peaks in Figure 4.3 have been attributed to static friction at the tops of strokes where the velocity is near zero [McDonald, Yasuhara]. My own experiments on the other hand indicate that the separate peaks are at least not completely a friction artifact. Subjects were asked to make an up-down finger movement on plexiglass. With this arrangement there is negligible friction between pen and surface and between hand and surface since the fingers only are moving off the surface. Sharp peaks are evident in the resultant acceleration recording (Figure 4.4).

4.2 Measurements of Handwriting

1. An Underlying Oscillation

The oscillatory nature of human handwriting is illustrated clearly by simple patterns such as chains of *e*'s and *u*'s. A typical example of such writing is presented in Figure 4.5. From the velocity traces it is evident that a constant frequency of oscillation is maintained through the extent of this writing. Whether the oscillation is sinusoidal, trapezoidal, or some other pattern is a question examined more closely in section 4.3.

The frequency of oscillation is a strong function of letter height. Generally speaking, smaller frequencies are associated with greater letter heights. For an alternating pattern of *e*'s and *l*'s, such as the example of Figure 4.6, the frequency difference between the *e* and *l* is typically around $\sqrt{2}$. A fair degree of rhythmic constancy is maintained among letters of the same height, i.e., the *e*'s are all made at the same frequency and

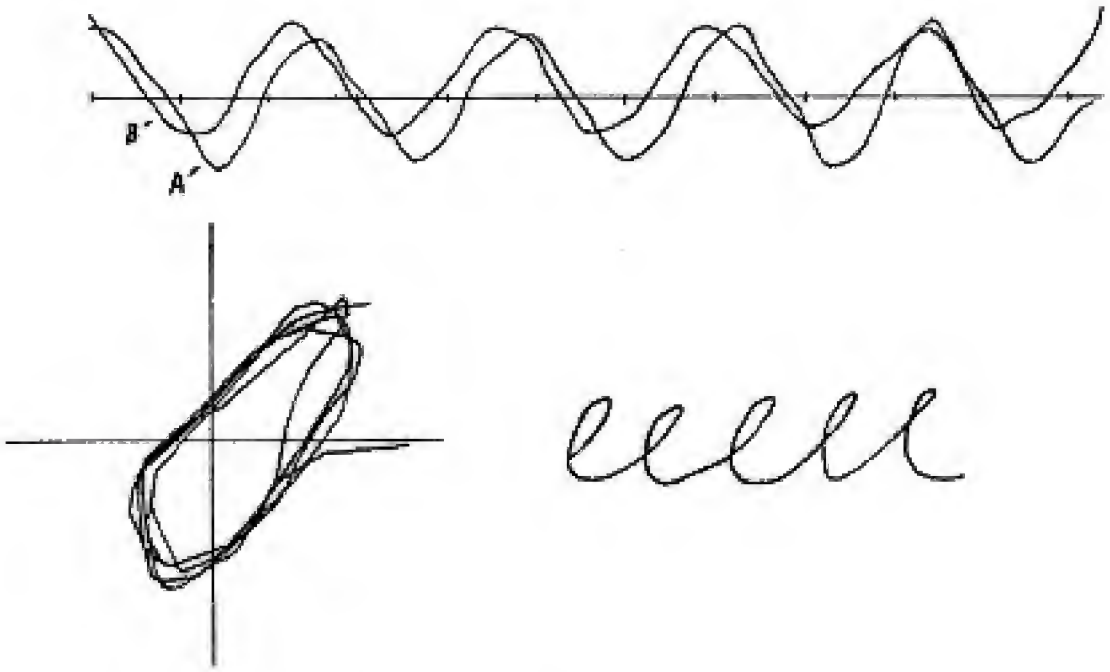


Figure 4.5. Recording of a string of *e*'s by subject STE. Top diagram: vertical (A) and horizontal (B) velocity traces. Bottom left: velocity space diagram.

the *l*'s are all made at the same frequency. The issue of height control is dealt with more thoroughly in a later section of this chapter.

2. Horizontal Modulation of Shape

Within a string of letters of the same height, the frequency of oscillation is observed to vary somewhat, although the variation is within bounds of frequencies of letters greater or smaller in height. The model of the previous chapter would lead one to expect that for letters of a given height a constant vertical oscillation frequency is maintained and a modulation of the horizontal oscillation acts independently to shape a letter. A reason for maintaining a constant vertical frequency, it was suggested, was ease of maintaining a straight base line.

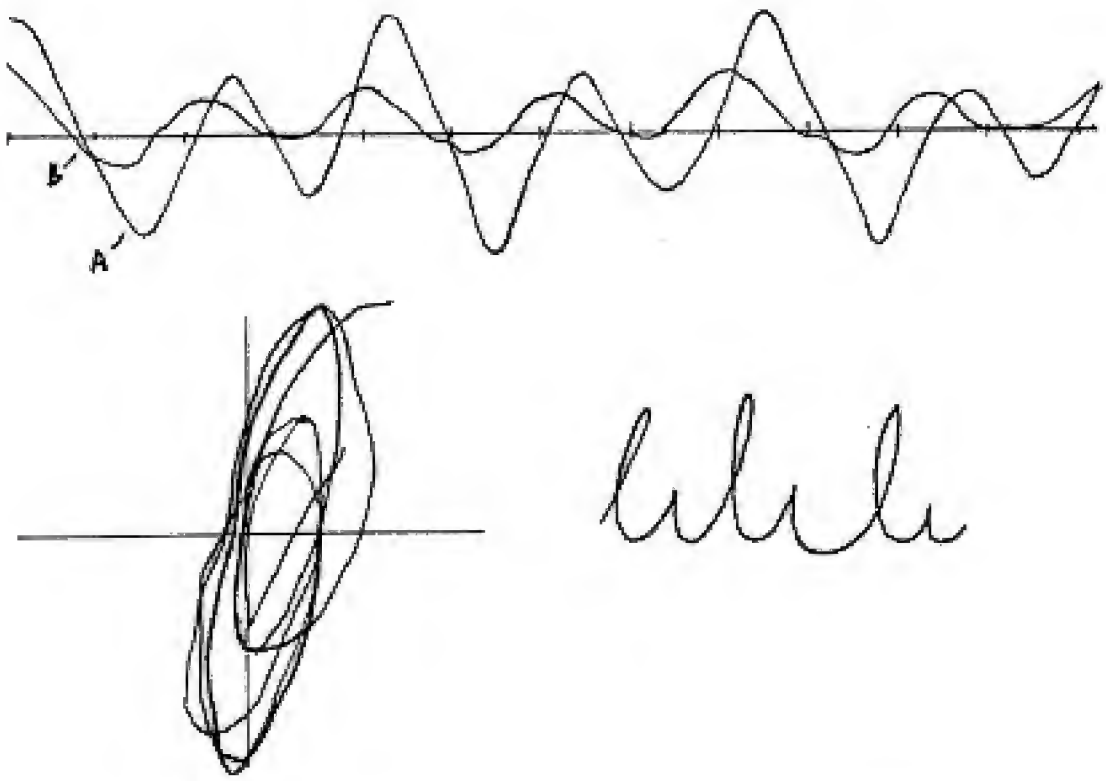


Figure 4.6. A recording of *lalala* by subject HOI, and the associated velocity space diagram shows constancy of slant, shape, and horizontal sweep velocity. Top diagram: vertical (A) and horizontal (B) velocity traces. Bottom left: velocity space diagram.

To illustrate frequency variation among different letters, two examples of signatures are presented in Figure 4.7. In Table 4.1 the frequencies and letter heights for individual down-up strokes is presented for each signature. Only down-up strokes that return to the same height are considered. In Figure 4.7A the strokes from the *a* to the *s* and the stroke *o* are roughly equal in height, and the frequencies range from 4.1 to 3.5. Particularly for the *o* there has been some modulation of vertical frequency, which shows that a vertical modulation participates with the horizontal modulation for letter shape. The down-up stroke for the *n* is half the height of the previous strokes and has a correspondingly higher frequency 4.9. Similarly the clockwise down-up stroke at the top of the *a* and the up-down stroke at the bottom of the *a*, both of very small height,

Figure 4.7. (A): Signature by subject LAR; (B) signature by subject STE.

are executed at the high frequencies 6.3 and 6.7 respectively. The second signature in Figure 4.7B shows a greater degree of rhythmic regularity than the previous example. The range of frequencies is tight, between 5.9 and 6.3. Curiously the height variation is greater than the frequency variation.

Table 4.1A		
Stroke	Frequency (Hz)	Height (mm)
top of *a*	6.3	0.4
a	4.1	2.7
a to *r*	4.1	2.6
r to *s*	3.8	3.1
o	3.5	2.6
n	4.9	1.2
bottom of *s*	6.7	0.4

Table 4.1B		
Stroke	Frequency (Hz)	Height (mm)
e to *v*	5.9	2.4
v	6.3	2.1
v to *e*	5.9	1.6
e to *n*	6.3	1.6
n	5.9	2.0
n to *s*	6.3	1.8

3. Constant Horizontal Velocity Sweep

In the mathematical treatment in Chapter 2 it was assumed that the writing movement could be factored into an oscillatory movement superimposed on a constant horizontal velocity sweep. This assumption was substantially validated by measurements on human writing. Subjects were asked to produce strings of letters *lelele*. The linear sweep for a given letter in the string was estimated by fitting a least squares ellipse to the velocity space diagram. The mean value for the constant velocity sweep c and the standard deviation $c(\sigma)$ for the letters within one of these strings was tabulated in Table 4.2 for a number of subjects and for 3 trials for each subject; the units are cm/sec.

Table 4.2						
Subject	c	$c(\sigma)$	β	$\beta(\sigma)$	*zero*	*zero*(σ)
BRU	9.5	3.4	78.2	4.6	-10.8	6.5
	13.6	2.2	70.2	7.9	-17.8	10.3
	12.9	2.1	76.0	7.0	-19.6	4.2
GRI	6.0	1.1	63.6	9.7	-10.8	4.7
	7.5	3.8	65.0	10.2	-12.5	3.6
	9.4	2.1	63.8	6.0	-18.0	4.6
HOL	11.5	1.5	74.5	0.5	-5.5	2.5
	12.0	3.4	76.3	5.8	-12.7	6.6
	13.3	2.4	79.3	2.3	-8.7	3.7
LAR	12.4	3.6	62.6	6.8	-12.0	1.7
	12.7	2.5	67.1	2.4	-8.9	4.3
	16.0	4.5	58.8	2.3	-9.2	2.6
MAS	12.3	2.3	64.0	8.4	-11.2	4.0
	12.0	1.0	66.0	1.0	-7.0	0.0
	17.0	1.4	68.0	4.5	-5.3	4.8
MAT	28.0	6.4	63.7	2.9	-16.3	4.5
	26.3	3.6	64.0	1.6	-8.5	1.8
	30.2	4.9	63.8	1.5	-14.8	3.8
MCD	19.0	6.3	59.3	8.7	-16.5	9.2
	16.2	3.9	57.2	9.5	-29.0	9.3
	20.5	5.3	54.2	8.2	-27.3	9.8
SJO	10.5	1.3	77.8	3.8	-8.2	1.6
	16.6	1.6	78.6	3.9	-15.8	5.3
	16.8	2.5	76.4	4.8	-11.6	2.2
STA	11.3	3.0	55.5	6.2	-16.7	7.0
	11.3	1.3	54.5	4.2	-20.0	4.8
	10.8	2.9	56.0	3.3	-16.2	4.4
STE	12.5	2.6	52.3	6.0	-17.8	5.8
	6.5	4.8	69.0	4.6	-13.2	2.3
	11.8	3.3	66.0	3.6	-16.8	5.0
WOO	11.0	3.7	67.5	3.2	-5.7	2.2
	11.4	1.5	65.8	5.2	-6.8	6.4
	13.8	6.2	66.5	5.2	-6.2	3.5

The standard deviations show a fair degree of constancy of e from letter to letter within a given string. Some subjects showed a greater degree of constancy than others. The same observations apply between writing samples for a given subject. The standard deviations are actually less significant than might appear because

the horizontal velocity amplitudes, not indicated in Table 4.2, are typically around 3 times greater than the constant velocity sweep.

4. Height Control

The theoretical treatment in Chapter 2 examined two possible strategies for height control: modulation of amplitude and modulation of frequency. Actual experimental measurements indicate that both modulations take place. Subjects were asked to produce a 6 letter sequence of alternating *e*'s and *l*'s. Measuring each letter from bottom corner to bottom corner, the ratios of *l* to *e* letter duration, vertical velocity amplitude, and height were calculated and averaged over 3 trials (Table 4.3). The average over all subjects is also given.

Table 4.3

Subject	Duration ratio	Velocity ratio	Height ratio
BRU	1.39	1.75	2.25
GRI	1.25	1.97	2.46
HOL	1.39	2.11	2.92
LAR	1.39	1.99	2.59
MAS	1.52	2.05	2.71
MAT	1.40	1.81	2.36
SJO	1.34	2.06	2.52
STE	1.44	2.34	2.98
ULL	1.26	1.85	2.12
GROUP	1.38	1.99	2.55

The results show that vertical velocity amplitude is approximately double for the *l* as compared to the *e*. This doubled amplitude is obtained in roughly equal measure from a frequency modulation roughly proportional to $\sqrt{2}$ and an acceleration amplitude modulation roughly proportional to $\sqrt{2}$. Because of the double integration of acceleration the proportional contribution to height for frequency vs. amplitude modulation is $2\sqrt{2}$.

The frequency modulation is not a result of limitations in the power plant. Muscles involved in handwriting seem grossly overpowered for the task. There are two pieces of evidence for this assertion. (1) As mentioned earlier handwriting frequency was not affected when weights were attached to the accelerometer

apparatus. (2) Handwriting frequency is independent of writing size. This independence is an accepted observation in the handwriting literature [Denier van der Gon and Thuring 1965, Yasuhara 1975] and is an observation substantiated by my own measurements. Subjects were asked to produce the alternating *el* patterns in different sizes. The *l* height and the time required to produce one *el* pair were calculated and averaged over 3 trials (Table 4.4).

Table 4.4

Subject	*l* height (in)	*el* time (msec)
BRU	0.38 in	445
	0.86	435
MAT	0.34	547
	0.46	558
SJO	0.20	403
	0.31	423
STE	0.25	522
	0.37	503
	0.49	498

Table 4.4 indicates that the amount of time required to write an *el* pair is independent of size over a factor of 2 across a sampling of subjects.

In Figure 4.8 the vertical acceleration profile for two different sizes of the same word, one twice the size of the other, nearly overlap. Thus the handwriting muscles are capable of writing tall letters in the same time as short letters. Yet an *e* in large writing similar in shape but taller than an *l* in small writing is written in less time, in fact in precisely the same time as an *e* in the small writing.

The frequency modulation by a factor of $\sqrt{2}$ fits particularly well the alternate activation mode of the spring model. Presumably an extra amplitude modulation would be required because of stylistic constraints to produce a sufficiently tall *l*. This magnitude of frequency modulation is also well within the range of a variable spring constant model.

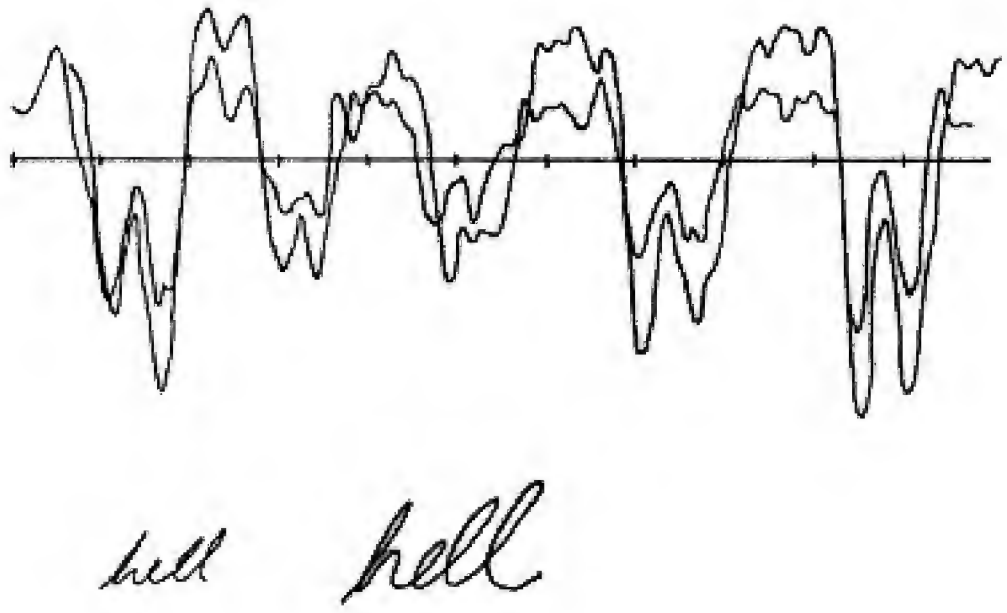


Figure 4.8. Vertical acceleration profiles from two different writing sizes of *hell*, one double the other, by subject STF show a high degree of burst duration agreement.

5. Slant Constancy

The extent to which subjects maintained constant slant during writing was determined by computing the slant of individual letters in the sequence *lelele*. The slant of a letter was computed by fitting a least squares ellipse to the velocity space diagram corresponding to that letter and by determining the slant from the coefficients of this ellipse. These measurements are model based in the sense that a sinusoidal oscillation is presumed to underly the experimental velocity space diagram. The results in Table 4.2 under the column β and $\beta(\sigma)$ are expressed as a mean value of slant and standard deviation for the letters in one sequence *lelele*.

Table 4.2 shows that there are mixed results with regard to consistency of writing slant. For consistency of slant within a single word or letter sequence, some subjects during certain trials had highly consistent writing slants with a standard deviation below 4 degrees. At the other extreme some subjects exhibited poor slant

control with standard deviations above 8 degrees. Most subjects fell more or less between these extremes. For consistency of slant between trials, the results are also mixed. Generally the range of slant means between trials did not exceed 10 degrees, with some subjects having a range of 5 degrees or less. Finally, although this breakdown is not given in Table 4.2, the slants of *e*'s were more or less the same as the slants of neighboring *l*'s, with a tendency in some subjects to make the *e*'s more slanted.

It might be concluded that generally speaking people are moderately successful at maintaining a consistent writing slant.

6. Shape Constancy

It was suggested in Chapter 2 that shape control at top corners is related to the horizontal velocity at the top vertical velocity zero crossing. To what extent is shape constancy maintained in going between *e*'s and *l*'s? In Table 4.2 the mean horizontal velocity at the vertical velocity zero crossing, labeled "zero" in Table 4.2, and the associated standard deviation are presented for the sequence *lelele*. As opposed to the previous columns in Table 4.2, the zero values are direct measurements rather than extrapolated values from a least squares ellipse; the units are cm/sec.

A rough shape constancy is demonstrated by Table 4.2. Thus subjects seem to attempt to keep the same top shape when going between *e*'s and *l*'s.

4.3 Discussion of Results

The Forcing Function

A major advantage of the spring muscle model is that a sinusoidal oscillation can be obtained with minimal effort. Starting from an appropriate set of initial conditions, a sinusoidal oscillation propagates indefinitely with quite passive control: i.e., the zero settings are not changed. The main conclusions of Chapter 2 however are not dependent on the validity of the spring muscle model. A sinusoidal forcing function arising from active central programming would also be subject to the same slant, shape, and height constraints as a sinusoidal

function arising from the spring model. Only the details of the modulations would differ.

In terms of the position or velocity recordings for writing of a given height, the fit of the spring model and of the various flavors of step patterns is about equally good. Actual velocity space diagrams are ambiguous with regard to being ellipsoidal or rounded parallelogram. Figure 4.9 shows best fit ellipse and a best fit rounded parallelogram to Figure 4.5 and the associated synthetic writing. Both agree fairly well with the data. Less ambiguous is the diagram in Figure 4.10A, harboring a parallelogram hidden by a nearly collapsed diagram. Shifting the horizontal velocity, which does not change the nature of the velocity space diagram, brings out the underlying parallelogram (Figure 4.10B).

With regard to recorded accelerations, trapezoidal and sinusoidal patterns seem to fit the data for fast writing about equally well. With progressively slower writing, however, large acceleration plateaus emerge (Figure 4.11). Assuming the underlying control mechanism does not change with writing speed, an assumption which seems born out by the similarity of the acceleration profiles, the fast writing bursts would seem to have a trapezoidal basis.

The sharp negative acceleration peaks, present in Figure 4.11 and in most other writing data, are foreign to both patterns, however. As indicated in section 4.1, these peaks are not just friction artifacts. They may indicate a segmentation of the acceleration profiles at the top corners: although the underlying pattern is continuous, it might be thought of as a chain of down-up strokes. Some writing specimens may show more than two peaks per burst, however, and may show peaks in the positive bursts as well.

Two Joint Horizontal Movement

A mechanical writing configuration of two independent joints, one executing an up-down movement and the other a left-right movement, has been a frequent assumption in past handwriting research [Eden 1962, Memelstein 1964, Denier van der Gon and Thuring 1965, Koster and Vredenbregt 1971, Yasuhara 1975]. Up-down movements have usually been ascribed to the fingers, back-forth movements to the wrist. I have observed in subjects other configurations that yield an essentially orthogonal two joint movement, involving

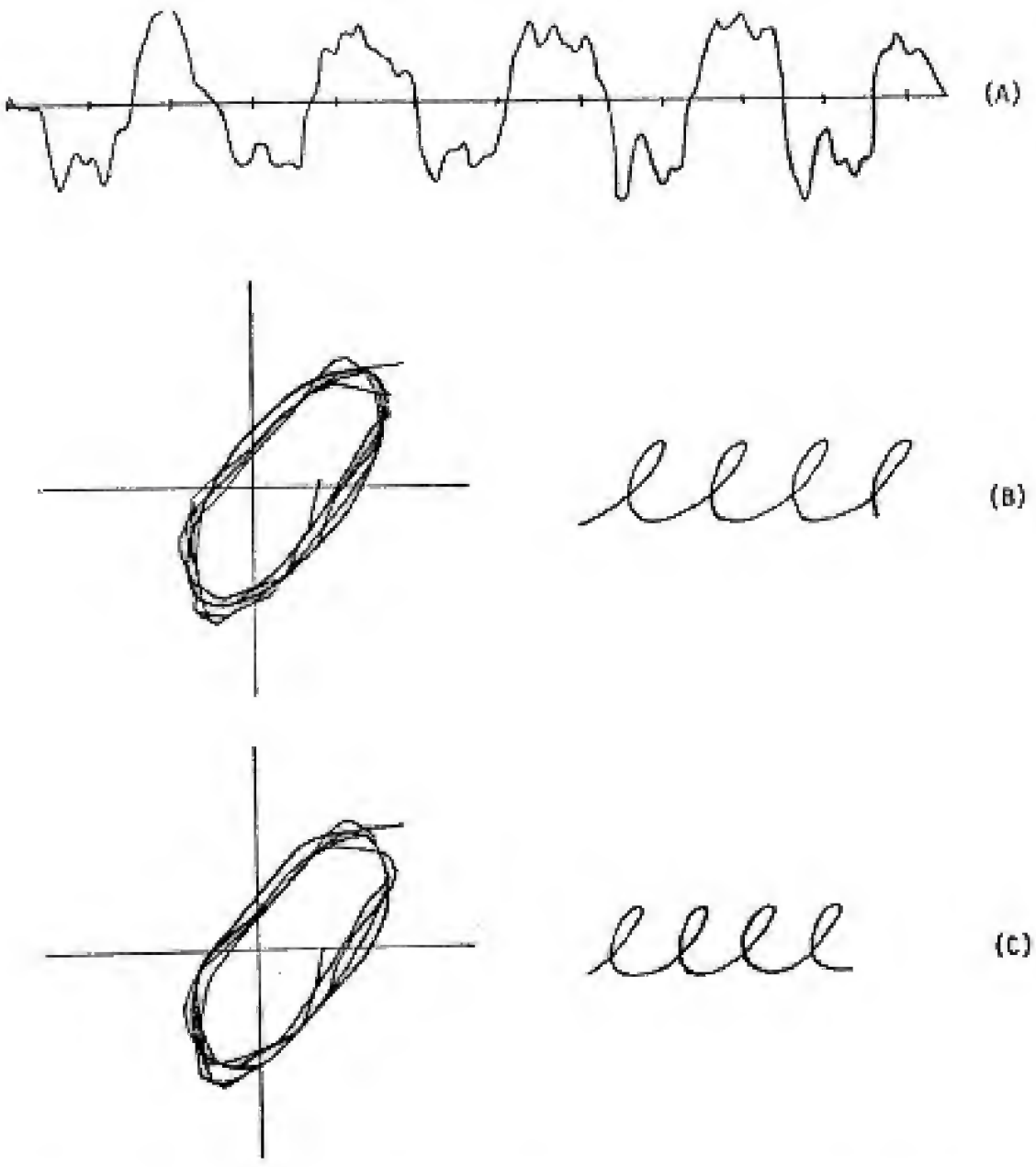


Figure 4.9. (A) The vertical acceleration corresponding to Figure 4.5 is more trapezoidal than sinusoidal. (B) A best fit ellipse and the resultant writing. (C) A best fit rounded parallelogram and the resultant writing.

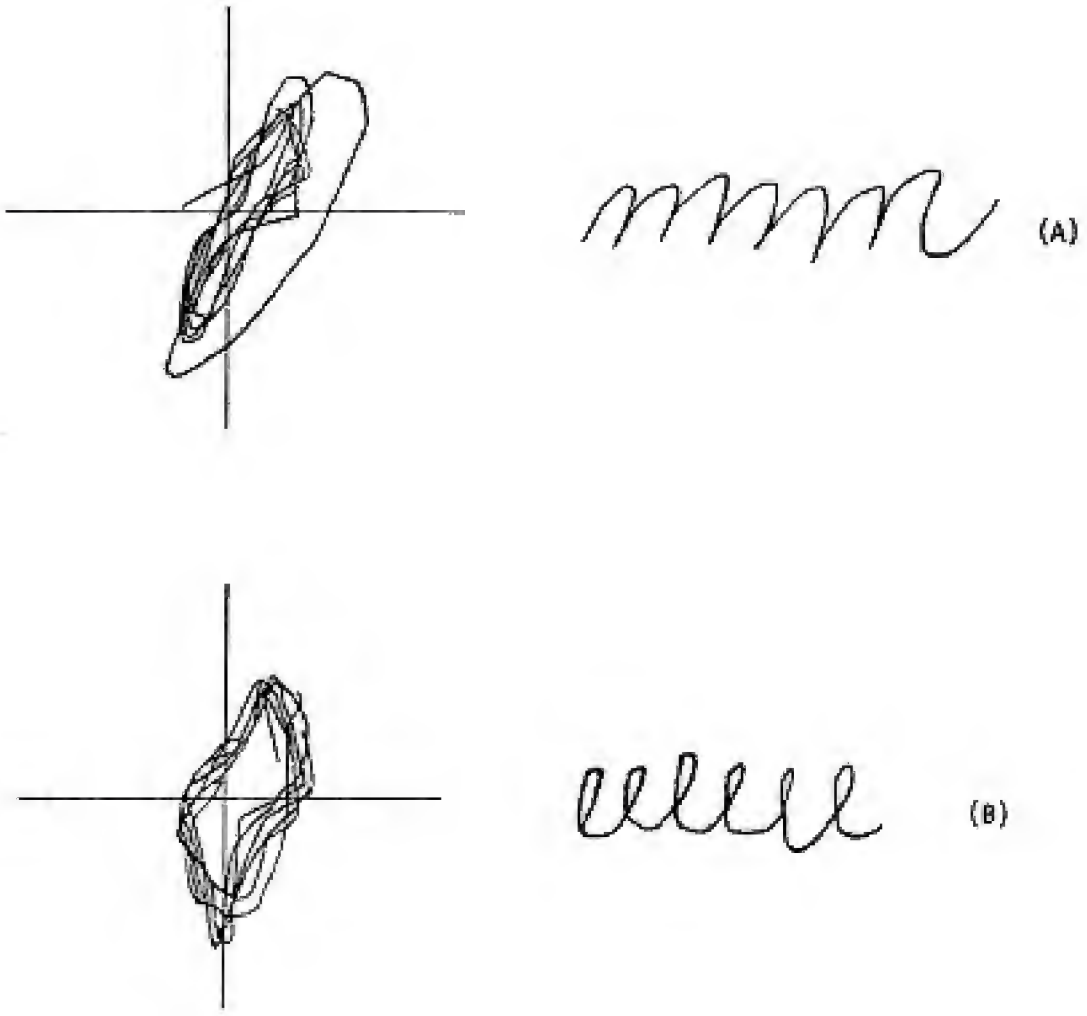


Figure 4.18. By phase shifting the horizontal velocity pattern in (A), the parallelogram nature of the velocity space diagram (B) emerges.

the elbow and the shoulder. The method of using the fingers and wrist also varies depending on grasp and degree of pronation. Lefthanders switch the roles of fingers and wrist. Each subject most likely hit upon their own particular configuration, insofar as formal handwriting instruction normally concentrates only on letter shape and not on the mechanism of writing [Hertzberg 1926, Hagan 1976].

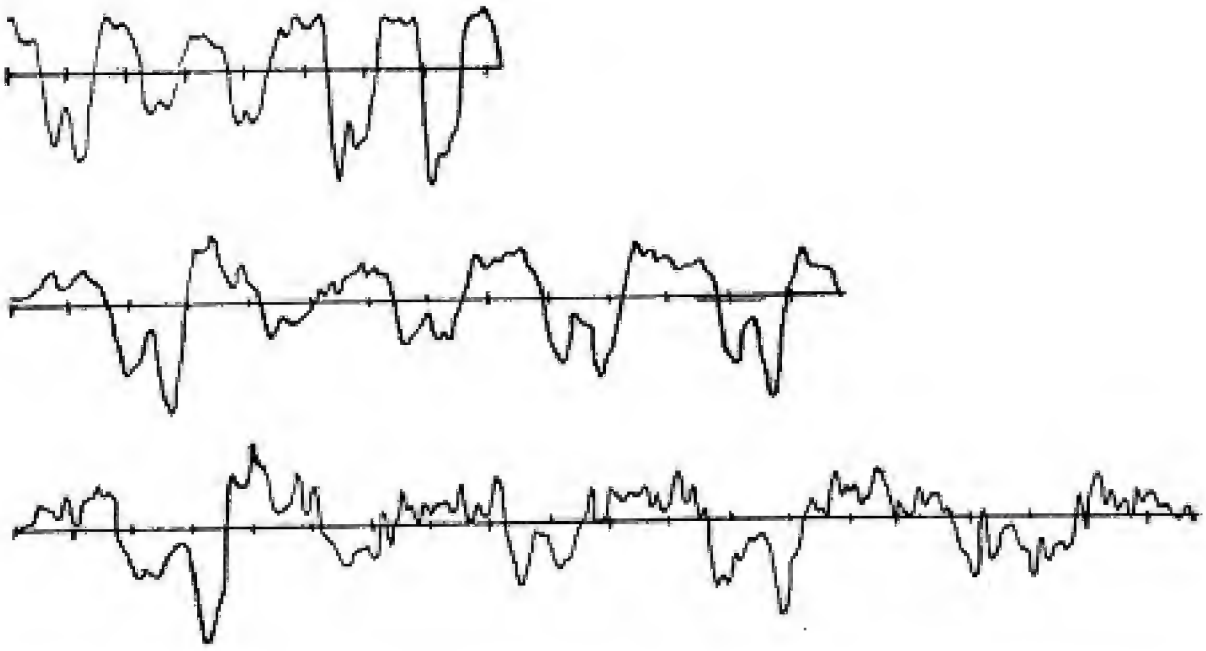


Figure 4.11. Vertical acceleration profiles from three samples of *hall* written at different speeds show trapezoidal bursts.

The disposition of the two joints need not be orthogonal. A joint disposition test can be devised by asking the subject to produce repetitive down-up and then back-forth movements; the angle between the two traces is often similar to the writing slant and may indicate the joint disposition. One might be tempted to say of subjects who evidenced a slanted down-up movement similar to their writing slant that the "vertical" joint accelerates along this slanted direction.

The agreement found between writing slant and the angle produced during up-down movements by certain subjects is also open to suspicion. Using just their fingers, these same subjects could produce any desired writing slant. This supports the notion that handwriting might actually be a three-joint movement: two joints act as oscillatory x-y generators, a third joint produces a constant horizontal sweep. Subjects find it possible to draw circles with fingers alone, so that there are at least two degrees of freedom here. This may be coupled

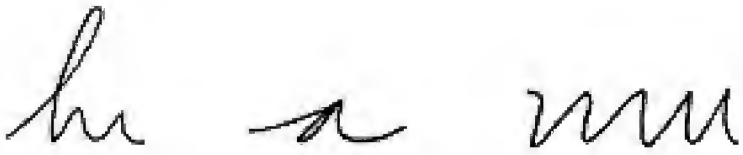
with a horizontal sweep at the wrist or at the elbow. In any case what one might be observing with diagonaling up-down movements is simultaneous activation of the two finger degrees of freedom; the resultant line is in the direction of the writing slant.

More careful studies of the joints and muscles used in handwriting than heretofore conducted are required to clear up the issue of writing slant. In the absence of such a study in this thesis, all joints are assumed orthogonal unless there are compelling reasons to assume the contrary. In any case the vertical record is unaffected save for a constant scaling factor, and one is safe in making arguments based on the vertical record.

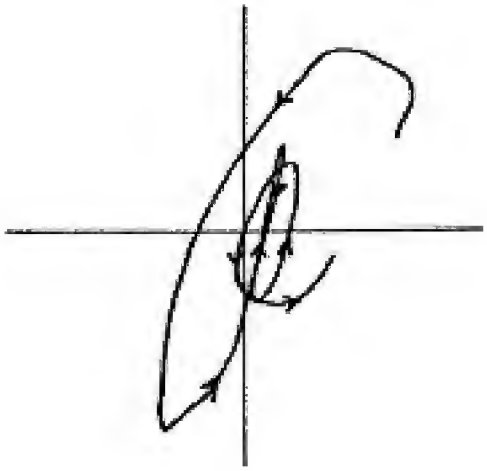
A decomposition of the horizontal movement into a linear component and a periodic component makes the idea of a two-joint horizontal movement attractive. The larger more proximal joint would be responsible for the linear movement, the smaller distal joint for the periodic movement. The forearm with wrist or fingers and the wrist with fingers are two pairs that may satisfy the proximal-distal joint arrangement. Indeed, one of the subjects was taught to write with a constant forearm sweep. In any case assertions in the literature that handwriting is a two joint movement needs more careful study than has heretofore been given. The advantage of the three joint movement would be to simplify the programming of the horizontal movement.

Clockwise elimination

There are a number of modifications and shortcuts taken by fast writers to standard cursive script shapes in order to adapt the shapes to requirements of speed and rhythmicity. In Figure 4.12 the *h* in *hi*, the letter *a*, and the *m* in *mu* are examples of modified shapes produced by three different fast writing subjects. The common feature of these letters in the Palmer script is a clockwise movement: to round the top of the *a*, and to produce the bottom cusp corners of *h* and *m*. In examining the velocity space diagrams for these letters by fast writing subjects, for example that of *hi* in Figure 4.12B, what stands out is that clockwise movement has been completely eliminated in favor of a uniform counterclockwise movement. At most a straight line velocity space diagram is obtained, corresponding to zero phase difference between horizontal and vertical joints (Figure 2.21J); that is to say, clockwise shapes are approximated with an essentially rounded sawtooth pattern.



(A)



(B)

Figure 4.12. (A) The letters *h* in *hi*, *a*, and *m* in *mm* are examples of modified letter forms by three different fast writers. (B) The velocity space diagram for the *hi* shows only counterclockwise movement.

Chapter 5. The Minimum Energy Movement

Though considerable effort has been expended in the study of the human motor system, the execution of even simple movements is not well understood. One current theory holds that movements are memorized in terms of final position [Bizzi et al.]. The final position theory has as a basis the spring muscle model discussed in Chapter 3. Referring to Figure 3.2, the final position theory maintains that the position L_1 can be reached independent of starting position merely by setting rates n_1 and g_1. This theory is interesting from a manipulation viewpoint because it obviates the need for precise trajectory calculation in the case of single joint movement.

There are many choices of agonist-antagonist length-tension curve pairs that have L_1 as equilibrium position. One choice that could be expected to require less energy is n_2 and g_2, which minimize the isometric tensions. More generally, it is conceivable that some complex sequence of innervation rates n_i, g_i might require less energy than a scheme which selects the final length-tension curves immediately. The determinaton of this optimal innervation pattern is the focus of this chapter.

5.1 Spring Muscle Model Solution

The investigated properties of muscle present a too complicated view for analytic treatment. The plan here is to simplify the muscle mechanics until an analytic solution to the optimal energy problem is possible, then to examine if the nature of the solution is changed by adding some of the excluded muscle properties.

Referring to the spring muscle model of Figure 3.3 and associated equation (3.1), define a control variable U and a state variable X as below.

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x \\ \dot{x} \end{bmatrix} \qquad U = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} x_0 \\ x_n \end{bmatrix} \tag{5.1}$$

Setting the mass $m = 1$, the state variable representation of the spring system is:

$$\dot{X} = \begin{bmatrix} 0 & 1 \\ -k_0 - k_m & -b \end{bmatrix} X + \begin{bmatrix} 0 & 0 \\ k_0 & k_m \end{bmatrix} U \tag{5.2}$$

More compactly,

$$\begin{aligned} \dot{X} &= AX + BU \\ &= f(X, U, t) \end{aligned} \tag{5.3}$$

Muscle Energetics

The energy E expended during movement equals work plus heat. The work W may be subdivided into conservative work performed on the mass m and nonconservative work performed on the viscous element. The isometric heat Q_t is given off in maintaining the muscle at a particular tension P_0. The rate of energy expenditure is thus:

$$\begin{aligned} \dot{E} &= P_0 v \text{ (power)} + a P_0 \text{ (maintenance heat rate)} \\ &= (x_2 + a)(k_0(u_1 - x_1) + k_m(x_1 - u_2)) \end{aligned} \tag{5.4}$$

where v is velocity and a is the maintenance heat coefficient. The two force terms have been summed because each contributes to energy loss. We have excluded the shortening heat because the active damping was also excluded, and because there may be a theoretical relationship between the two [Huxley, Caplan]. The transient characteristics of heat production have also been excluded.

The Euler-Lagronge Equations

The task now is to find the time varying control $U(t)$ that minimizes the energy used in moving between two points in a fixed interval of time. Let V represent the energy consumed in applying the control U to yield the trajectory X. The problem of minimizing V is readily approached by techniques of modern control theory. The fundamental equations that the optimal control $U(t)$ must satisfy are derived from a theorem from the *calculus of variations*. This theorem states that in order to find the n-vector $X(t)$ that minimizes $V(X)$, where

$$V(X) = \int_0^\infty L(X, \dot{X}, t)dt \tag{5.5}$$

subject to the constraint relations

$$g_i(X, \dot{X}, t) = 0 \qquad i = 1, \ldots, m \leq n \tag{5.6}$$

then $X(t)$ satisfies the Euler equations

$$\frac{\partial L'(X, \dot{X}, t)}{\partial x_i} - \frac{d}{dt}\frac{\partial L'(X, \dot{X}, t)}{\partial \dot{x}_i} = 0 \qquad i = 1, \ldots, n \tag{5.7}$$

where

$$L'(X, \dot{X}, t) = L(X, \dot{X}, t) + \sum_{i=0}^{m} \lambda_i(t) g_i(X, \dot{X}, t) \tag{5.8}$$

and $\lambda_i(t)$'s are the multiplier functions [Schultz and Melsa].

Applying this theorem to the optimal control problem, the state equations $\dot{X} = f(X, U, t)$ represent the equality constraints. L represents the rate of change of energy $\dot{E}$. The Hamiltonian $H = L + \lambda^T f$ represents (5.8), where $\lambda^T = [\lambda_1 \ \lambda_2]$. By applying the Euler equation first for X and then for U, it can be shown that the minimizing $U(t)$ satisfies the following two Euler-Lagrange equations [Schultz and Melsa].

$$\dot{\lambda} = -H_X \tag{5.9}$$

$$H_U = 0 \tag{5.10}$$

The Minimum Principle

Because L is linear in the control U, there will not generally exist a minimum energy solution. To obtain a realistic solution, constraints must be placed on the control. The solution in this case will lie on the constraint

boundaries [Bryson and Ho]. Constraints on U, however, make it impossible to differentiate H with respect to U.

The minimum principle of Pontryagin makes it possible to proceed from this point. Pontryagin showed that even if the control is constrained, one still obtains a minimal solution by finding the $u^0 = u^0(X, \lambda, t)$ to minimize the Hamiltonian H, but by inspection rather than by differentiation. After finding the minimizing u^0, one forms $H^0 = H(X, u^0, \lambda, t)$ and then solves the following two equations [Schultz and Melsa].

$$\dot{X} = \frac{\partial H^0}{\partial \lambda} \tag{5.11}$$

$$\dot{\lambda} = -\frac{\partial H^0}{\partial X} \tag{5.12}$$

There are two natural constraints that fall on the control U. First, the spring cannot push.

$$u_1 - x_1 \geq 0 \tag{5.13}$$

$$x_1 - u_2 \geq 0 \tag{5.14}$$

Second, springs have a maximum tension that they can exert. Without this constraint the solution would involve an infinite impulse. For the moment we assume the maximum tension is constant and independent of length:

$$u_1 - x_1 \leq c_1 \tag{5.15}$$

$$x_1 - u_2 \leq c_2 \tag{5.16}$$

where c_1 and c_2 are constants. The case of maximum tension varying with length is deferred until section 5.2.

A Bang-Coast-Bang Solution

To facilitate inspection of the Hamiltonian, we expand $H = \lambda^T f + L$ into three lines, the first depending on u_1, the second on u_2, and the third on neither control.

$$H = k_y(u_1 - x_1)(a + x_2 + \lambda_2) + k_n(x_1 - u_2)(a + x_2 - \lambda_2) + x_2(\lambda_1 - b\lambda_2) \tag{5.17}$$

To minimize H with respect to u_1, we observe that if $a + x_2 + \lambda_2 > 0$ then H is minimized when $u_1 = x_1$. If $a + x_2 + \lambda_2 < 0$ then H is minimized with $u_1 = x_1 + c_1$. Similarly, it can be shown for u_2 that when $a + x_2 - \lambda_2 < 0$ the minimizing u_2 lies at $x_1 - c_2$; otherwise u_2 is at x_1. Combining these results, one finds a bang-coast-bang solution to the minimum energy for muscle movement.

Case 1: $\lambda_2 < -(a + x_2)$.
Then $u_1 = x_1 + c_1$, $u_2 = x_1$.
Case 2: $|\lambda_2| \leq (a + x_2)$.
Then $u_1 = x_1$, $u_2 = x_1$.
Case 3: $\lambda_2 > (a + x_2)$.
Then $u_1 = x_1$, $u_2 = x_1 - c_2$.

The Solution Equations

Substituting the minimizing u^0 into H, one obtains three functions corresponding to the three cases.

Case 1: $H^0 = k_y c_1(a + x_2 + \lambda_2) + x_2(\lambda_1 - b\lambda_2)$
Case 2: $H^0 = x_2(\lambda_1 - b\lambda_2)$
Case 3: $H^0 = k_n c_2(a + x_2 - \lambda_2) + x_2(\lambda_1 - b\lambda_2)$

The differential equation (5.11) and its solution becomes for the three cases:

$$\dot{X} = \begin{bmatrix} 0 & 1 \\ 0 & -b \end{bmatrix} X + \begin{bmatrix} 0 \\ kc \end{bmatrix}$$

$$x_2(t) = x_2(t_0)e^{-b(t-t_0)} + \frac{kc}{b}(1 - e^{-b(t-t_0)}) \tag{5.18}$$

$$x_1(t) = x_1(t_0) + \frac{1}{b}(x_2(t_0) - x_2(t)) + \frac{kc}{b}(t - t_0) \tag{5.19}$$

where $c = c_1$ and $k = k_y$ for case 1; $c = 0$ for case 2; and $c = -c_2$ and $k = k_n$ for case 3. The time t_0 represents the starting time. The differential equation (5.12) and its solution are:

$$\dot{\lambda} = \begin{bmatrix} 0 & 0 \\ -1 & b \end{bmatrix}\lambda - \begin{bmatrix} 0 \\ kc \end{bmatrix}$$

$$\lambda_1(t) = \lambda_1(t_0) \tag{5.20}$$

$$\lambda_2(t) = \lambda_2(t_0)e^{b(t-t_0)} + \frac{\lambda_1(t_0) + kc}{b}(1 - e^{b(t-t_0)}) \tag{5.21}$$

where c and k have the same meaning as above except $c = +e_2$ for case 3. Since $\lambda_1(t)$ is constant, it appears henceforth as λ_1 without a time dependence.

The Extremal Versus Singular Solution

It is proven in appendix A that there are exactly three events in the extremal bang-coast-bang solution: an acceleration period, a coast period, and a deceleration period. No other combination of bangs and coasts is minimizing. However, a nonextremal minimizing solution may arise from a singular arc at the switching points. The Hamiltonian (5.17) has the curious property that if $\lambda_2 = |\alpha + x_2|$ then the corresponding control may take on any value and still minimize H. If a control can be found to maintain $\lambda_2 = |\alpha + x_2|$ for a finite time interval, then a non-extremal solution to energy minimization might exist. This situation is called a singular arc and arises from a performance index linear in control but quadratic in state [Bryson and Ho]. To maintain $\lambda_2 = |\alpha + x_2|$ for a finite time interval, all time derivatives of the two switching curves must be zero:

$$\frac{d^n(\lambda_2 + \alpha + x_2)}{dt^n} = 0 \qquad n \geq 0 \tag{5.22}$$

$$\frac{d^n(\lambda_2 - \alpha - x_2)}{dt^n} = 0 \qquad n \geq 0 \tag{5.23}$$

Carrying through the analysis for a singular arc at the first switching point (5.22), the time varying force during the singular arc is:

$$k_y(u_1 - x_1) = \lambda_1 + b\alpha + 3bx_2(t_1)e^{2b(t-t_1)} + \frac{3}{2}(b\alpha + \lambda_1)(e^{2b(t-t_1)} - 1) \tag{5.24}$$

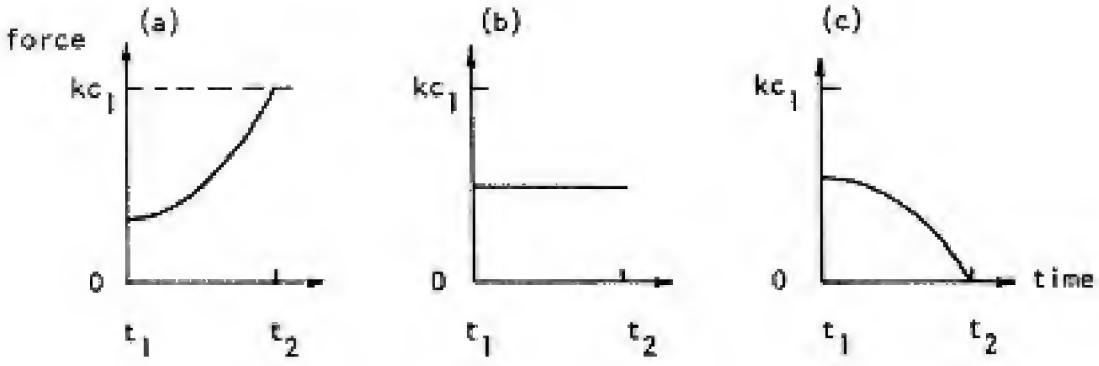


Figure 5.1. The three possible forms for force in a singular arc solution.

Unfortunately no sufficient condition has yet been developed to test whether a singular arc is minimizing, and one must compare values of the performance index for specific parameter values for the singular arc solution versus the extremal solution. Depending on the choice of λ_1 the force (5.24) takes one of the three forms in Figure 5.1.

Of these forms only 4C has been found minimizing for some parameter combinations. To search for such combinations, a set of parameters was initially deduced from Rack and Westbury (Table 5.1). The elapsed distance x_f and the elapsed time t_f are variable and have been chosen as 0.2 cm and 0.4 sec respectively. The initial and final velocities are assumed zero. For the extremal solution bang,coast,bang there result 8 nonlinear equations in 8 unknowns from (5.18)-(5.21) and the initial conditions (x_f, t_f).

Table 5.1

Parameter	Value
k	2.0 kg/cm
m	0.2 kg (plausible value)
b/m	3.16 /sec (chosen to give $\zeta = 0.5$)
c	1.0 cm
a	0.1 cm/sec (deduced from Woledge)

For the singular solution bang,4C,coast,bang 15 nonlinear equations in 15 unknowns result from (5.18)-(5.23) and the initial conditions. The equations were solved numerically by Newton-Raphson and gradient methods. Individual parameters were varied and energies of movement computed from (5.4). Solving (5.4), the energy for the extremal solution is

$$E = kc(\Delta x_1(t_1) + \Delta x_1(t_3)) + kc\alpha(\Delta t_1 + \Delta t_3) \tag{5.25}$$

where t_1 is the switching time from acceleration to coast, t_3 is the time at the end of deceleration, Δt_1 is the duration of acceleration, $\Delta x_1(t_1)$ is the distance moved during acceleration, and $\Delta x_1(t_3)$ is the distance moved during deceleration. For the singular solution, the energy is

$$E = kc(\Delta x_1(t_1) + \Delta x_1(t_4)) + kc\alpha(\Delta t_1 + \Delta t_4) + k_v(u_1 - x_1)(x_2(t) + \alpha) \tag{5.26}$$

where t_1 is the switching time from acceleration to the singular arc 4C, t_2 is the switching time from 4C to coast, and t_4 is the time at the end of deceleration. The force $k_v(u_1 - x_1)$ is given by (5.24), while the velocity $x_2(t)$ is:

$$x_2(t) = x_2(t_1)e^{2b(t-t_1)} + \frac{b\alpha + \lambda_1}{2b}(e^{2b(t-t_1)} - 1) \tag{5.27}$$

The energies for the extremal versus the singular solution are compared in Tables 5.2a-g; the units are kg cm/kg wt. In Table 5.2c the parameters k and b are varied simultaneously but at a fixed damping ratio of 0.5. In Tables 5.2f-g the parameters k and b are respectively set at 16 and 4 rather than at the Table 5.1 values where the extremal solution is minimizing over the whole range of x_f and α. The initial values from Table 5.1 are starred in Table 5.2. A singular solution becomes minimizing with high values of k, b, c, and t_f, and with low values of α and x_f. As the parameters cause the coast time to approach zero (higher b and x_f, lower k, c, and t_f), the singular and extremal solutions become identical because the 4C portion vanishes.

For the extremal solution it is proved in appendix B that there is an upper limit on the duration of coast. It is tempting to speculate that for longer coast durations a singular solution becomes minimizing, but the singular solution in Table 5.2 is not always minimizing under these conditions. Perhaps a different combination of bangs, coasts, and singular arcs would then be minimizing, but this remains an open question. Some combinations can be proved impossible, such as bang, 4C, coast, 4A-C, bang.

Table 5.2a

k	Singular	Extremal
7	*****	0.708
8	0.66117	0.66106
*10	0.6173	0.6164
20	0.5640	0.5627
30	0.5513	0.5509
40	0.54554	0.54552
50	0.5423	0.5426
100	0.536	0.537

Table 5.2b

b	Singular	Extremal
2.3	*****	0.545
2.4	0.552036	0.552034
2.6	0.5675	0.5674
3.0	0.6020	0.6014
4.0	0.7045	0.7030
5.0	0.820	0.824
10.0	1.45	1.51
15.0	2.13248	2.13246
15.557	*****	2.4

Table 5.2c($b = k$)

k	Singular	Extremal
7	*****	0.664
8	0.632026	0.632023
*10	0.6173	0.6164
14	0.6442	0.6440
20	0.708	0.725

Table 5.2d

c	Singular	Extremal
0.7	*****	0.708
0.8	0.6612	0.6611
*1.0	0.6173	0.6164
2.0	0.5640	0.5627
3.0	0.5513	0.5508
4.0	0.54554	0.54551
5.0	0.5427	0.5437
6.0	0.5401	0.5407

Table 5.2e		
t_f	Singular	Extremal
0.35	*****	0.814
0.36	0.76123	0.76122
*0.40	0.6173	0.6164
0.45	0.512	0.511
0.50	0.444	0.445
0.60	0.358	0.369

Table 5.2f($k = 16, b = 4$)		
x_f	Singular	Extremal
0.006	*****	0.0041
0.050	0.0632	0.0634
0.100	0.1925	0.1942
*0.200	0.6647	0.6681
0.300	1.4636	1.4620
0.350	2.0120	2.0072
0.400	2.6878	2.6825
0.540	*****	6.4598

Table 5.2g($k = 16, b = 4$)		
a	Singular	Extremal
0.06	0.6112	0.6161
*0.10	0.6647	0.6681
0.20	0.7974	0.7981
0.25	0.8633	0.8631
0.50	1.1903	1.1882
1.00	1.8393	1.8382
1.70	2.748417	2.748411
1.80	*****	2.8784

5.2 Spring Model Relaxations

A natural question is whether the minimum energy solution is changed by incorporating a more realistic muscle model. For those relaxations of the spring model involving only X dependencies, the answer is that the solution remains bang-coast-bang. The reason is that the Hamiltonian H remains linear in the control U, and the minimization of H with respect to U occurs at fixed X. Whether the solution also remains acceleration-coast-deceleration needs to be determined for each case.

Relaxations of the spring model involving X dependencies include the following.

1. Position Limits on Tension

For real muscle the maximum isometric tension varies with position (Figure 4.1). This makes c_1 and c_2 into functions of x_1, but the controls will still fall at the extremes wherever they are.

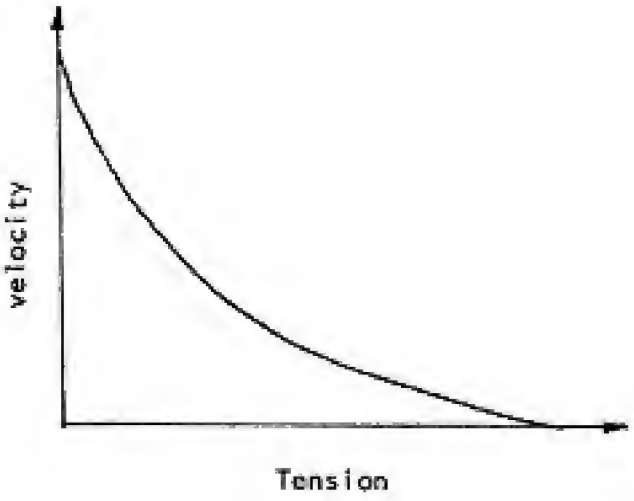


Figure 5.2. Tension dependence on velocity (Hill 1938).

2. Velocity Limits on Tension

Actual muscle exhibits a hyperbolic force-velocity relation. If P_0 is the isometric tension, then the maximum force P that can be produced for a velocity v is [Hill 1938] (see Figure 5.2):

$$P = P_0 - \frac{v(P_0 + a)}{v + b'} \tag{5.28}$$

The term $(P_0 + a)/(v + b')$ can be considered the coefficient of active damping. The coefficient a has been determined as $.25P_0$; the force P then becomes $P_0 - 1.25P_0v/(v + b')$. The literature conflicts on the value of active damping during lengthening. For consistency with the shortening heat (below), it is assumed the same as active damping during shortening.

Associated with the active damping is an extra heat expenditure above the isometric heat due to shortening. This shortening heat rate is [Hill 1964]:

$$\dot{Q}_s = .16P_0v + .18Pv \tag{5.29}$$

The isometric heat rate $\dot{Q}_i$ remains aP_0, but the power is now Pv. Substituting the sum of spring forces for the isometric tension P_0 and (5.29) for P, the energy rate is:

$$\dot{L} = (k_s(u_1 - x_1) + k_n(u_2 - x_1))(a - 0.09x_2 + \frac{1.45b'}{x_2 + b'}) \tag{5.30}$$

Similarly it can be show that the equation of motion is:

$$\dot{x}_2 = -bx_2 + (1 - \frac{1.25x_2}{x_2 + b'})(k_s(u_1 - x_1) - k_n(x_1 - u_2)) \tag{5.31}$$

When these terms are combined to form the Hamiltonian, the control is seen to remain linear. Hence the solution is once again bang-coast-bang.

3. Spring Constant Variations with Position

One way of bringing the simplified length-tension curves of Figure 4.2 closer to those of Figure 4.1 is illustrated in Figure 5.3. The spring constant k_n varies with position, but at any given position the constant k_n is the same for all controls u_2. Under these conditions the solution remains bang-coast-bang.

4. Parallel and Series Elastic Elements

The incorporation of these elements into the model is depicted in Figure 5.4. Since the parallel elastic element depends only on position, it does not change the solution. The series elastic elements and the active springs may be replaced with equivalent springs with constants $k_s' = k_s k_s/(k_s + k_s)$ and $k_n' = k_n k_s/(k_n + k_s)$. This modification also has no effect on the solution.

5. Stiffness Regulation

With regard to U dependencies, in Figure 3.1 the spring constant k_n is seen to vary with firing rate at any fixed position. The linear portions of these length-tension curves when extended seem to intersect at a

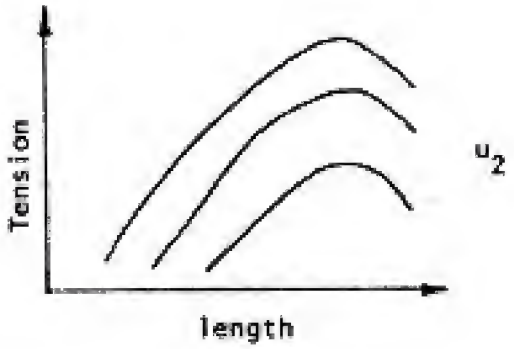


Figure 5.3. Hypothetical length-tension curves with the porperty that at any given lenght the slopes are the same for all choices of u_2.

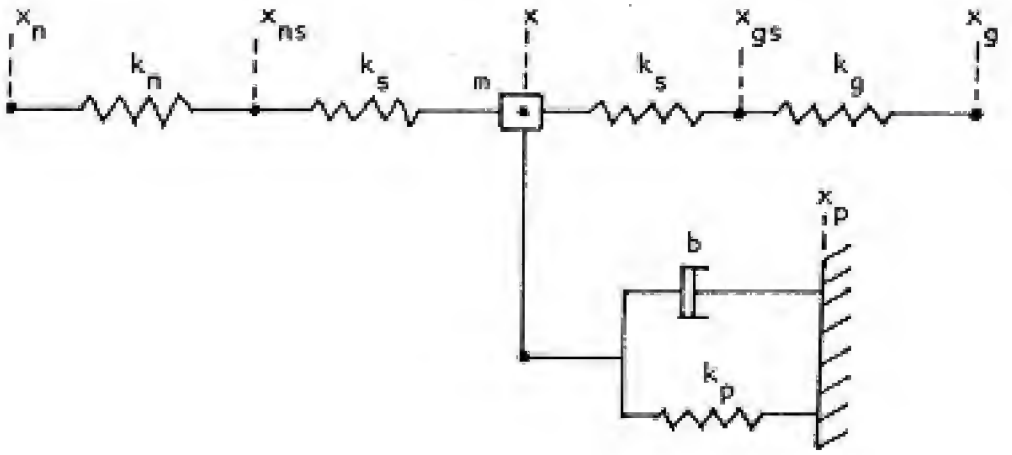


Figure 5.4. An expanded muscle model incorporating series elastic elements k_s and a parallel elastic element k_p.

common point (Figure 3.8).

In this circumstance the spring constant k_n is $T_0/(u_2 - x_0)$. The u_2 terms of the Hamiltonian H become:

$$H' = \frac{T_0}{u_2 - x_0}(x_1 - u_2)(\alpha + x_2 - \lambda_2)$$
$$= T_0(\alpha + x_2 - \lambda_2)(\frac{x_1 - x_0}{u_2 - x_0} - 1) \tag{5.32}$$

If $\alpha + x_2 - \lambda_2 > 0$ then H' is minimized at $u_2 = x_1$; otherwise $u_2 = x_1 - c_2$. That is to say, the solution for u_2 is exactly the same as in section V. A similar analysis holds for u_1. Thus the minimizing pattern is also bang-coast-bang.

Chapter 6. Concluding Remarks

The oscillation theory of handwriting suggests that letter shapes emerge as individuations of an underlying oscillation. The alternative view which is dispelled by this theory is that each letter has a separate motor program which can be invoked to produce that letter in isolation, and that the word formation process is one of linking together the motor programs for the desired letters. In the oscillation theory there is a preexisting and underlying repeated pattern of letter shapes, for example a cycloidal chain of *e*'s for a sinusoidal based oscillation, and that this pattern propagates indefinitely unless it is modulated. Rather than an active process of forming letter shapes, there already exist letter shapes typical of the oscillation pattern and the modulations serve to remold the preexisting letter shapes into the desired letters. A modulation will change the underlying oscillation pattern to a new one, which like the old will propagate indefinitely unless it too is modulated. In a sinusoidal based oscillation, for example, an original *e* cycloid can be modulated to an *l* cycloid, and after this modulation the new underlying pattern is the *l* cycloid.

In motor control work one is used to thinking in terms of motor programs, and here the motor programs are best considered the sequence of modulations. The underlying oscillatory process acts as an interpretive program that "interprets" the motor programs, which are the sequence of modulations, in the context of the current oscillation.

In creating a word a temporal sequence of modulations to some oscillation pattern must be set up. Even for writing single isolated letters such as an *a* an oscillation must be created and one or two modulations applied to it. There is a question as to the size and the nature of the conceptual unit in handwriting. Is the conceptual unit linguistically based, such as a syllable or a word, or is it based on some oscillation feature, such as whether the movement is clockwise or counterclockwise or whether adjacent letters have the same height? For typing it has been suggested that the conceptual unit is the word [refs], and the finger sequence to produce a word is considered as a single motor program. What the conceptual unit in handwriting is and whether it too should be considered a single motor program remains an open question.

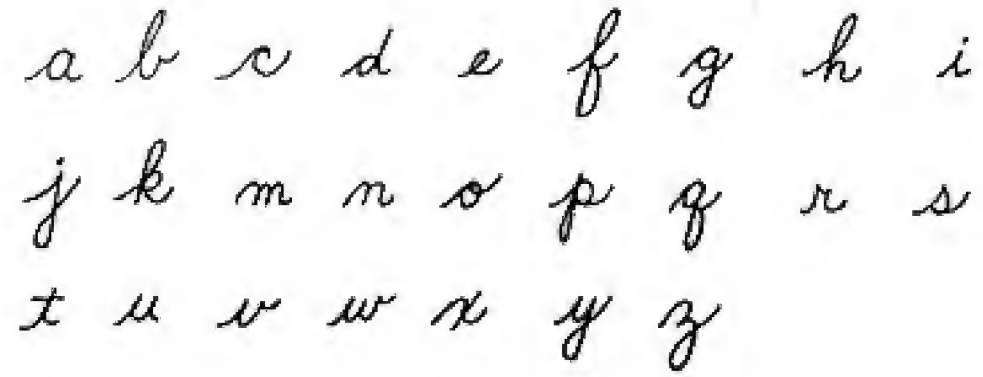
The human motor system has configured itself to make the handwriting act a relatively simple task. It has factored the large numbers of degrees of freedom of hand and arm into just a few degrees of freedom that facilitate the control of handwriting. Coupled oscillations in x and y directions produce diverse corner shapes; a potential third degree of freedom for a horizontal constant velocity movement separates the corners to form letters and words. The process of letter shaping reduces largely to controlling the vertical velocity zero crossing in the velocity space diagram: the intercept controls corner shape while the slope at the zero crossing controls writing slant. Under the constraints of the zero crossings letter height is modulated by both frequency and amplitude modulation of the acceleration.

The oscillation-modulation scheme reduces the information processing requirements for handwriting at the expense perhaps of letter shape diversity. It might be speculated that a reduction of the information processing complexity for handwriting is necessary for thinking and writing at the same time.

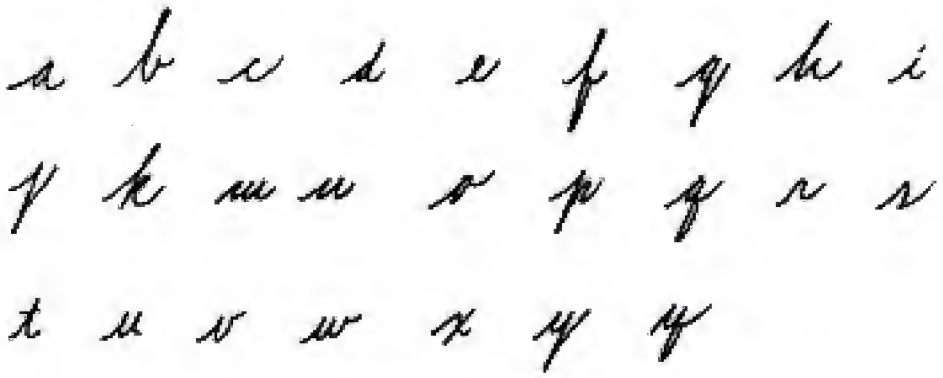
Applications of This Research

The reason for eliminating clockwise movements may be to reduce the number and degree of modulation, as discussed in Chapter 4. It would seem that a new cursive script needs to be devised that takes into account the ease of character formation with speed. The elimination of clockwise movement is a first step. Streamlining other shapes to eliminate excess curvature or excessive loop width would limit the amount of phase shift. A neutral rhythm, namely a sequence of top cusp shapes, could then be set up that is easily modulated for other shapes; a small phase modulation in one direction would produce loops, in the other direction rounded top corners.

One possibility for a new script is given in Figure 6.1B, where the proposed script is compared with the Palmer script of Figure 6.1A. The clockwise top roundings of the Palmer letters *a*, *c*, *d*, *g*, *o*, and *q* have been eliminated in favor of a pattern which is essentially a closed *u*. The clockwise bottom loops of letters *g*, *j*, and *y* have been virtually eliminated in favor of a more angular bottom; it should be considerably easier to follow these letters with an *e*, the difficulty of which was discussed in Chapter 3. The clockwise bottom cusp



(A)



(B)

Figure 6.1. (A) Standard Palmer cursive script; (B) A proposed new script that eliminates clockwise movement.

corners of *h*, *k*, *m*, and *n* have been replaced with a sort of angular corner produced when there is zero phase shift between joints. The letter *k* is somewhat problematical in a counterclockwise scheme. If the second downstroke more or less follows the previous upstroke, a letter form distinct from *h* can be produced. Letters *p*, *s*, and *z* would also required substantial modification. One possibility for *z* is to relate it to *y* in the same

way that q is related to g.

One might speculate that the neuronal structures responsible for the oscillation and modulation components of handwriting are separately identifiable. Handwriting has been used as a tool for diagnosing neurological diseases, and it may be possible to make a more specific diagnosis by observing exactly how the handwriting has degenerated.

Suggestions for Further Study

The most pressing problem requiring further study is the clarification of joint roles. More detailed observations than heretofore obtained of the hand during writing are needed to decide such issues as whether a third degree of freedom is executing a linear movement and whether the oscillatory horizontal and vertical degrees of freedom are actually nonorthogonal. The role of the thumb requires clarification. The integration of downward pressure of the pen on the writing surface with the planar aspect of movement is also necessary.

The nature of the underlying oscillation also needs further clarification, such as the extent to which the oscillation is mechanically based or actively programmed. The nature of the forcing function is also unclear, whether it is sinusoidal, rectangular, trapezoidal, or something else. The acceleration recordings show an asymmetry between positive and negative veritical acceleration bursts. The negative bursts tend to show a separation into two peaks; the negative burst amplitudes show greater variation than the positive burst amplitudes. This asymmetry is in need of explanation.

Lastly, this thesis has concerned itself only with the formation of lower case script in the Palmer style of English writing. The extent to which ideas in this thesis transfer to the production of upper case Palmer script, printing, or scripts for other languages such as Arabic, Hebrew, and Chinese, remains an open question.

References

Bizzi, Emilio, Andres Polit, and Pietro Morasso [1976]: Mechanisms underlying achievement of final head position, **J. Neurophys.**, 39:435-444.

Bryson, A.E. and Yu-Chi Ho [1969]: **Applied Optimal Control.** Ginn and Company, Waltham, Mass.

Caplan, S.R. [1966]: A characteristic of self-regulated linear energy converters. The Hill force-velocity relation for muscle. **J. Theor. Biol.**, 11:63.

Collins, Carter C., David O'Meara and Alan B. Scott [1975]: Muscle tension during unrestrained eye movements. **J. Physiol.** 245:351-369.

Cook, G. and L. Stark [1967]: Derivation of a model for the human eye positioning mechanism. **B. Math. Biophys.** 29:153-174.

Crane, H.D. and R.E. Savoie [1977]: An on-line data entry system for hand-printed characters. **Computer**, March 1977, 43-50.

Denier van der Gon, J.J., and Thuring, J.Ph. [1965]: The Guiding of Human Writing Movements. **Kybernetik**, 2, 145-148.

Eden, M. [1962]: Handwriting and Pattern Recognition. **IRE Trans. on Information Theory**, IT-8, 160-166.

Feldman, A.G. [1974a]: Change of muscle length due to shift of the equilibrium point of the muscle-load system. **Biofizika** 19:534-538.

Feldman, A.G. [1974b]: Control of muscle length. **Biofizika** 19:749-753.

Grillner, S. [1975]: Locomotion in Vertebrates: Central Mechanisms and Reflex Interaction. **Physiol. Rev.**, 55, 247-304.

Hagan, Patti [1976]: Johnny not only can't write; he can't handwrite, either. **New York Times Sunday Magazine**, May 23, P. 64.

Halliday, A.M. and Redfearn, J.W.T. [1956]: An Analysis of the Frequencies of Finger Tremor in Healthy Subjects. **J. Physiol. Lond.**, 134, 600-611.

Hennemann, E. [1974]: Peripheral Mechanisms Involved in the Control of Muscle. Chapter 22 in V.B. Mountcastle, **Medical Physiology**, vol. 1, 617-635.

Herbst, N.M. and C.N. Liu [1977]: Automatic signature verification based on accelerometry. **IBM J. Res. Develop.**, May 1977, 245-253.

Hertzberg, O.E. [1926]: A Comparative Study of Different Methods Used in Teaching Beginners to Write. No. 214, Teachers College, Columbia University.

Hill, A.V. [1938]: The heat of shortening and the dynamic constants of muscle. **Proc. R. Soc. Lond. (Biol.)**, 126:136.

Hill, A.V. [1964]: The effect of load on the heat of shortening of muscle. **Proc. R. Soc. Lond. (Biol.)**, 159:297.

Hill, A.V. [1970]: **First and Last Experiments in Muscle Mechanics**. Cambridge University Press, London.

Huxley, A.F. [1957]: Muscle structure and theories of contraction. **Prog. Biophys. Chem.**, 7:255.

Koster, W.G., and Vredenbregt, J. [1971]: Analysis and Synthesis of Handwriting. **Medicine and Sport, vol. 6: Biomechanics II**. pp. 77-82.

Lawrence, J. Dennis [1972]: **A Catalog of Special Plane Curves**. Dover Publications.

McDonald, J.S. [1966]: Experimental Studies of Handwriting Signals. **RLE Technical Report 443**, MIT.

Mermelstein, P. [1964]: Computer Recognition of Connected Handwritten Words. PhD Thesis, Massachusetts Institute of Technology, January 1964.

Mermelstein, P. and M. Eden [1964]: Experiments on Computer Recognition of Connected Handwritten Words. **Information and Control**, 7:255-270.

Paul, R. [1972]: Modelling Trajectory Calculation and Servoing of a Computer Controlled Arm. Stanford Artificial Intelligence Memo 72.

Pellionisz, A. and R. Llinas [1979]: Brain Modeling by Tensor Network Theory and Computer Simulation. The Cerebellum: Distributed Processor for Predictive Coordination. **Neuroscience**, 4:324-348.

Rack, P.M.H. and D.R. Westbury [1969]: The effects of length and stimulus rate on tension in the isometric cat soleus muscle. **J. Physiol.**, 204:443-460.

Raibert, Marc H. [1977]: Motor Control and Learning by the State Space Model. AI-TR-439, MIT Artificial Intelligence Laboratory.

Schultz, D.G. and James L. Melsa [1967]: **State Functions and Linear Control Systems**. McGraw-Hill.

Stiles, R.N. and Randall, J.E. [1967]: Mechanical Factors in Hand Tremor Frequency. **J. Appl. Physiol. 23**, 324-330.

Thomas, George B. [1972]: **Calculus and Analytic Geometry**. Addison-Wesley.

Woledge, R.C. [1971]: Heat production and chemical change in muscle. In: **Progr. Biophys. Mol. Biol.**, 22:37.

Yasuhara, Makoto [1975]: Experimental Studies of Handwriting Process. **Rep. Univ. Electro-Comm. 25-2**, (Sci. and Tech. Sect.), 233-254.

Zierler, K.L. [1974]: Mechanism of muscle contraction and its energetics. In: **Medical Physiology**, V.B. Mountcastle, ed., 1:636-650.

Appendix A

In this appendix it is shown there are only two bangs and one coast in the extremal solution: one acceleration, followed by one coast period, terminated by one deceleration. No other combinations of coasts and bangs are possible. To demonstrate this is the only possible combination, it is necessary to examine the switching curves and their time derivatives.

The first lemma shows that once the control has passed from acceleration to coast, then the control cannot return to another acceleration but must proceed to deceleration. The second lemma shows that once deceleration has started, the deceleration must continue until the end of the movement. This proves that the acceleration-coast-deceleration combination is the only possible one. In the following it is presumed that the movement starts with acceleration in the positive x_1 direction. Hence all velocities are positive.

Lemma 1: After acceleration, the glide period cannot double back to another acceleration.

Proof: The proof of this lemma proceeds by examining the time derivative of the acceleration-coast switching curve (henceforth referred to as the slope of the switching curve). The slope of this curve is initially positive at the transition from acceleration to coast. In order for another acceleration to follow the coast period, this slope must become negative, leading to a contradiction.

At the first switching time t_1 the acceleration-coast switching curve is zero.

$$\lambda_2(t_1) + a + x_2(t_1) = 0 \tag{A1}$$

After the acceleration period, the coast equations are:

$$\lambda_2(t) = \lambda_2(t_1)e^{b(t-t_1)} + \frac{\lambda_1}{b}(1 - e^{b(t-t_1)}) \tag{A2}$$

$$x_2(t) = -(\lambda_2(t_1) + a)e^{-b(t-t_1)} \tag{A3}$$

Thus

$$\lambda_2(t) + \alpha + z_2(t) = \lambda_2(t_1)e^{b(t-t_1)} + \frac{\lambda_1}{b}(1 - e^{b(t-t_1)}) + \alpha - (\lambda_2(t_1) + \alpha)e^{-b(t-t_1)} \qquad (A4)$$

The slope of this switching curve is:

$$e^{b(t-t_1)}(b\lambda_2(t_1) - \lambda_1) + b(\lambda_2(t_1) + \alpha)e^{-b(t-t_1)} \qquad (A5)$$

At $t = t_1$, the slope of the switching curve is:

$$2b\lambda_2(t_1) - \lambda_1 + b\alpha > 0 \qquad (A6)$$

One can show this quantity cannot be less than zero. Next, suppose the coast doubles back to another acceleration. At some point the slope must go through zero. This time t is found from (A5) as:

$$e^{2b(t-t_1)} = \frac{b(\lambda_2(t_1) + \alpha)}{\lambda_1 - b\lambda_2(t_1)} > 1 \qquad (A7)$$

Case 1: $\lambda_1 - b\lambda_2(t_1) > 0$.
Then $\lambda_2(t_1) + \alpha > 0$, contradicting (A1).
Case 2: $\lambda_1 - b\lambda_2(t_1) < 0$.
Crossmultiplying (A7) and collecting terms,
$2b\lambda_2(t_1) - \lambda_1 + b\alpha < 0$, contradicting (A6).

Thus after acceleration, the coast period must eventually arrive at the deceleration switching point.

Lemma 2: The movement is locked in deceleration until the end.

Proof: It will be shown that if deceleration ever switches to coast, then the slope of the coast-deceleration switching curve requires an immediate return to deceleration. Hence the movement is locked in deceleration until the end.

Suppose there is a time t_3 when deceleration switches to coast. At this point the coast-deceleration switching curve is zero.

$$\lambda_2(t_3) - a - x_2(t_3) = 0 \tag{A8}$$

The coast switching curve $\lambda_2(t) - a - x_2(t)$ is:

$$\lambda_2(t_3)e^{b(t-t_3)} + \frac{\lambda_1}{b}(1 - e^{b(t-t_3)}) - a - x_2(t_3)e^{-b(t-t_3)} \tag{A9}$$

The slope of (A9) is:

$$e^{b(t-t_3)}(b\lambda_2(t_3) - \lambda_1) + bx_2(t_3)e^{-b(t-t_3)} \tag{A10}$$

At time t_3 the slope (A10) is $b\lambda_2(t_3) - \lambda_1 + bx_2(t_3)$. This is positive since $\lambda_2(t_3) > 0$, $x_2(t_3) > 0$, and $\lambda_1 < 0$ (lemma 3). This means that deceleration would bounce off the coast boundary and immediately continue the deceleration. Furthermore, since the slope is positive, the deceleration would not immediately switch back to coasting, causing chattering.

Lemma 3: $\lambda_1 < 0$.

Proof: At the second switching point t_2 we have

$$\lambda_2(t_2) - a - x_2(t_2) = 0 \tag{A11}$$

From (A2) and (A3), this becomes

$$\lambda_2(t_1)e^{b(t_2-t_1)} + \frac{\lambda_1}{b}(1 - e^{b(t_2-t_1)}) - a + (\lambda_2(t_1) + a)e^{-b(t_2-t_1)} = 0 \tag{A12}$$

Rearranging,

$$\frac{\lambda_1}{b}(1 - e^{b(t_2-t_1)}) = -\lambda_2(t_1)[e^{b(t_2-t_1)} + e^{-b(t_2-t_1)}] + a(1 - e^{-b(t_2-t_1)}) \tag{A13}$$

From (A1) and (A3) we find an expression for $\lambda_2(t)$.

$$\lambda_2(t_1) = -\alpha - \frac{k_g c_1}{b}(1 - e^{-b(t_1 - t_0)}) \tag{A14}$$

Substituting into (A13),

$$\frac{\lambda_1}{b}(1 - e^{b(t_2 - t_1)}) = \frac{k_g c_1}{b}(1 - e^{-b(t_1 - t_0)})(e^{b(t_2 - t_1)} + e^{-b(t_2 - t_1)})) + \alpha(1 + e^{b(t_2 - t_1)}) \tag{A15}$$

Thus

$$\lambda_1 = \frac{k_g c_1(1 - e^{-b(t_1 - t_0)})(e^{b(t_2 - t_1)} + e^{-b(t_2 - t_1)}) + b\alpha(1 + e^{b(t_2 - t_1)})}{1 - e^{b(t_2 - t_1)}} \tag{A16}$$

Since the numerator is positive and the denominator is negative, λ_1 is negative.

Taken together, these lemmas show that acceleration passes through coast to deceleration. There is no possible variation in this scheme. It is also possible to show the movement cannot start by coasting followed by acceleration.

Appendix B

A surprising limitation on the value of $t_2 - t_1$, the duration of the coasting time, has been found. The switching curve during acceleration is:

$$\lambda_2(t) + a + x_2(t) = \lambda_2(t_0)e^{b(t-t_0)} + \frac{\lambda_1 + k_y c_1}{b}(1 - e^{b(t-t_0)}) + a + \frac{k_y c_1}{b}(1 - e^{-b(t-t_0)}) \tag{B1}$$

The slope of this switching curve is:

$$e^{b(t-t_0)}(b\lambda_2(t_0) - \lambda_1 - k_y c_1) + k_y c_1 e^{-b(t-t_0)} \tag{B2}$$

At the first switching time t_1, the switching function (B1) is zero. Rearranging (B1) for $t = t_1$,

$$e^{b(t_1-t_0)}(\lambda_2(t_0) - \frac{\lambda_1 + k_y c_1}{b}) = -\frac{\lambda_1 + k_y c_1}{b}) - a - \frac{k_y c_1}{b}(1 - e^{-b(t_1-t_0)}) \tag{B3}$$

Substituting (B3) into (B2), the slope at t_1 is:

$$-2k_y c_1(1 - e^{-b(t_1-t_0)}) - \lambda_1 - ba \tag{B4}$$

Substituting for λ_1 from (A16),

$$\frac{k_y c_1(1 - e^{-b(t_1-t_0)})(2 + e^{-b(t_2-t_1)} - e^{b(t_2-t_1)}) + 2ba}{e^{b(t_2-t_1)} - 1} \tag{B5}$$

Since slope$(t_1) \geq 0$ and since the denominator is positive, so is the numerator.

$$k_y c_1(1 - e^{-b(t_1-t_0)})(2 + e^{-b(t_2-t_1)} - e^{b(t_2-t_1)}) + 2ba \geq 0 \tag{B6}$$

Because $2 + e^{-b(t_2-t_1)} - e^{b(t_2-t_1)}$ is a decreasing function of t_2, at some point (B6) becomes zero. Solving then for $e^{b(t_2-t_1)}$:

$$e^{b(t_2-t_1)} = \frac{b\alpha + \sqrt{(b\alpha + k_g c_1(1 - e^{-b(t_1-t_0)}))^2 + (k_g c_1(1 - e^{-b(t_1-t_0)}))^2}}{k_g c_1(1 - e^{-b(t_1-t_0)})} + 1 \qquad (B7)$$

As t_2 increases, t_1 will decrease. However, t_1 does not decrease enough to offset the effect of the t_2 increase. If $\alpha = 0$, (B7) reduces to

$$e^{b(t_2-t_1)} = 1 + \sqrt{2} \qquad (B7a)$$

Strangely, in this circumstance $t_2 - t_1$ depends only on b.